

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00, 15/12, C07K 14/00, 14/435		A1	(11) International Publication Number: WO 99/58660
			(43) International Publication Date: 18 November 1999 (18.11.99)
(21) International Application Number: PCT/US99/09847			
(22) International Filing Date: 6 May 1999 (06.05.99)			
(30) Priority Data:			
60/085,093	12 May 1998 (12.05.98)	US	
60/085,094	12 May 1998 (12.05.98)	US	
60/085,105	12 May 1998 (12.05.98)	US	
60/085,180	12 May 1998 (12.05.98)	US	
60/085,927	18 May 1998 (18.05.98)	US	
60/085,906	18 May 1998 (18.05.98)	US	
60/085,924	18 May 1998 (18.05.98)	US	
60/085,922	18 May 1998 (18.05.98)	US	
60/085,923	18 May 1998 (18.05.98)	US	
60/085,921	18 May 1998 (18.05.98)	US	
60/085,925	18 May 1998 (18.05.98)	US	
60/085,928	18 May 1998 (18.05.98)	US	
60/085,920	18 May 1998 (18.05.98)	US	
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832			
		(US). FLORENCE, Kimberly [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). SHI, Yang-gu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). WEI, Fing-Fei [CN/US]; 242 Gravatt Drive, Berkeley, CA 94705 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119-4321 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). ENDRESS, Gregory, A. [US/US]; 9729 Claggett Farm Drive, Potomac, MD 20854 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US).	
		(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).	
		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(54) Title: 97 HUMAN SECRETED PROTEINS			
(57) Abstract			
<p>The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

97 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human

the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

5

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and
10 polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

15

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original
20 environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

25 In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing
30 to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron.

- 5 In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's

solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions.

- 5 Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M
- 10 NaH_2PO_4 ; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

- Note that variations in the above conditions may be accomplished through the
- 15 inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above,
- 20 due to problems with compatibility.

- Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid
- 25 molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

- The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of
- 30 single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA

that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA
5 backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids
10 joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more
15 detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of
20 modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a
25 heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation,
30 iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

(See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth
5 Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting
10 activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the
15 present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

20 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

The translation product of this gene shares sequence homology with tag-7 which is thought to be important in tumor metastasis and is itself a secretory protein
25 (See, Kiselev SL, et al., J Biol Chem. 273:18633 (1998) and Genetika. 1996 May; 32(5): 621-628. (Russian)), and a family of peptidoglycan recognition proteins involved in the innate immune response to peptidoglycan in species as diverse as insects and humans (See, Kang, D. et.al., PNAS 95:10078 (1998)).

Preferred polypeptides of the invention comprise the following amino acid
30 sequence: WAGTQEPTGLPSTLSRSESWDH (SEQ ID NO: 211). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in keratinocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, dermatological disorders, especially skin cancers such as melanoma.

- 5 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the integumentary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skin, cancerous and
10 wounded tissues) or bodily fluids (e.g., sweat, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
15 111 as residues: Ser-25 to Ala-31, Gln-146 to Ser-151, His-231 to Asn-236.

- The tissue distribution in keratinocytes and homology to tag-7 indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of dermatological disorders, especially skin cancers like melanoma, and integumentary tumors (e.g., keratoses, Bowen's disease, basal cell carcinoma,
20 squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma). Tag-7 was discovered when gene expression was compared in a metastatic (VMR-Liv) neoplastic cell line and a related nonmetastatic (VMR-O) neoplastic cell line by means of the differential display method. A fragment of cDNA corresponding to the tag-7 gene, differentially expressed in the metastatic cell line,
25 was isolated. The full-length tag-7 cDNA was cloned and its nucleotide sequence was determined. The gene sequence claimed in this patent application has significant homology to tag-7 and on that basis is expected to share significant biological activities with tag-7. Such activities can be assayed as set forth herein and by assays known in the art.

- 30 Additionally, the homology to a conserved peptidoglycan recognition protein family involved in innate immunity, suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention

of various skin disorders including congenital disorders (e.g., nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), injuries and inflammation of the skin (e.g., wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (e.g., lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (e.g., cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and immunotherapy targets for the above listed tumors and tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1177 of SEQ ID NO:11, b is an integer of 15 to 1191, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of this gene shares weak sequence homology with FGF Receptor Ligand-2 which is thought to be important in activating FGF receptor in mediating cell proliferative functions.

Preferred polypeptides of the invention comprise the following amino acid sequence: EIIHNLPTSRMAARTKKKNDIINIKVPADCNTRMSYYYKGS
GKRGEMESWLVMSSWSILDFEFLEARPQLFNLVYTEHSTYSGRHYTRERGGF
MVFKNSYSQLLLKRKDSLCAFIQPMALNIIHVPMSKCIFFAQSGPSTFRSLW
WCPHPISKCQLGLYSSQIRDIPYLA (SEQ ID NO: 212).

- EIIHNLPTSRMAARTKKKNDIINIKVPADCNTRMS (SEQ ID NO: 213),
YYYKGSKGKRGEMESWLVMSSWSILDFEFLEARPQLF (SEQ ID NO: 214),
NLVYTEHSTYSGRHYTRERGGFMVFKNSYSQLLLR (SEQ ID NO: 215),
KDSLCAFIQPMALNIIHVPMSKCIFFAQSGPSTF (SEQ ID NO: 216), and/or
5 RSLWWCPHPISKCQLGLYSSQIRDIPYLA (SEQ ID NO: 217). Polynucleotides
encoding these polypeptides are also provided.

This gene is expressed primarily in neutrophils.

- Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, abnormal immune reactions or disorders. Similarly, polypeptides and
antibodies directed to these polypeptides are useful in providing immunological probes
for differential identification of the tissue(s) or cell type(s). For a number of disorders
of the above tissues or cells, particularly of the immune system tissue and connective
15 tissues, expression of this gene at significantly higher or lower levels may be
routinely detected in certain tissues or cell types (e.g., immune, cancerous and
wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid
and spinal fluid) or another tissue or cell sample taken from an individual having such
a disorder, relative to the standard gene expression level, i.e., the expression level in
20 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
112 as residues: Met-1 to Met-6.

- The tissue distribution and homology to FGF Receptor Ligand-2 indicates that
the protein products of this gene are useful for detection, treatment, and/or prevention
25 of immune disorders, especially those that are mediated by neutrophil functions. They
can be utilized in the treatment of neural and immune disorders, or to stimulate
proliferation of vertebrate cells, raise antibodies, and to screen for antagonists useful
for inhibiting tumor growth. Moreover, the expression of this gene product suggests a
role in regulating the proliferation, survival, differentiation, and/or activation of
30 hematopoietic cell lineages, including blood stem cells. This gene product may be
involved in the regulation of cytokine production, antigen presentation, or other

processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1237 of SEQ ID NO:12, b is an integer of 15 to 1251, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of this gene shares sequence homology with glycosyl transferase, which is thought to be important in glycosylation of proteins (See Genbank Accession No. g2996578).

This gene is expressed primarily in osteoclastoma cells, melanocytes, haemopoietic cells and colon tissue, and, to a lesser extent, in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the skin, blood, skeletal system and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic system, epithelium and skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, musculo-skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another
15 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 113 as residues: Glu-136 to Pro-141, Ala-221 to Ser-227, Asp-307 to Pro-312, Lys-
20 355 to Gly-361, Phe-449 to Pro-454.

The tissue distribution in musculo-skeletal and immune tissues, and the homology to glycosyl transferase protein, suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the haemopoietic, skeletal and epithelial systems, and cancers thereof,
25 as well as disorders associated with incorrect post-translational modification of proteins (i.e. glycosylation). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1720 of SEQ ID NO:13, b is an integer of 15 to 1734, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of this gene shares sequence homology with human pleckstrin protein (See Genbank Accession No. g35518), which is thought to be important in platelet formation or activity. Therefore, it is likely that this gene also has activity in platelets.

This gene is expressed primarily in keratinocytes, and, to a lesser extent, in spleen and bone marrow.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions which include, but are not limited to, immune and clotting disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and blood clotting systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, blood clotting, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 114 as residues: Leu-38 to Gly-49, Lys-75 to Thr-80.

The tissue distribution in keratinocytes, spleen and bone marrow, and the homology to pleckstrin suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and/or treatment of immune system and

clotting disorders. Furthermore, since this protein is 50% identical to the Pleckstrin protein, it is an excellent candidate for a protein kinase C substrate. Identification of this protein as a target of protein kinase C, and the exploration of its role in protein kinase C mediated responses, such as inflammation, may lead to a better

5 understanding of the inflammatory response. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets
10 for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1526 of SEQ ID NO:14, b is an integer of 15 to 1540, where both a and b correspond to the positions of nucleotide
20 residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

The gene encoding the disclosed cDNA is thought to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in
25 linkage analysis for chromosome 17.

This gene is expressed primarily in infant liver/spleen tissues, T cells, bone marrow stromal cells, and thymus tissue, and, to a lesser extent, in brain and tonsils tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various immune system disorders and/or diseases. Similarly,

polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 115 as residues: Ser-46 to Arg-54.

The tissue distribution in liver/spleen tissues, T-cells, bone marrow stromal cells, and thymus tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of cancers, most notably cancers of the immune system. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in a variety of cells of the immune system suggests that this gene may be a player in the progression of these diseases, and may be a beneficial target for inhibitors as therapeutics. Furthermore, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia, since stromal cells are important in the production of cells of hematopoietic lineages.

25 The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1544 of SEQ ID NO:15, b is an integer of 15 to 1558, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The translation product of this gene shares sequence homology with angiopoietin-2, an anti-angiogenic factor. See, for example, Maisonpierre, et al., Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis.
20 Science. (1997) 277(5322): 55-60, incorporated herein by reference in its entirety. Based on the sequence similarity, the translation product of this gene is expected to share certain biological activities with Angiopoietin-2 as may be assessed by assays known in the art and described herein.

Preferred polypeptides of the invention comprise the following amino acid
25 sequence:

MFTIKLLLFIVPLVISSRIDQDNSSFDSLSPKSRFAMLDDVKILANGLLQLGH
GLKDFVHKTKGQINDIFQKLNIFDQSFYDLSLQTSEIKEEEKELRRTTYKLQVK
NEEVKNMSLELNSKLESLLKILLQQKVYLEEQLTNLIQNQPETPEHPEVTS
LKTFVEKQDNSIKDLLQTVEDQYKQLNQHSQIKEIENQLRRTSIQEPTEISLSS
30 KPRAPRTTPFLQLNEIRNVKHDGIPAECTTIYNRGEHTSGMYAIRPSNSQVFHV
YCDVISGSPWTLIQHRIDGSQNFNETWENYKYGFGRLDGEFWLGLEKIYSIVK
QSNYVLRIELEDWKDNKHIEYSFYLGNHETNYTLHLVAITGNVPNAIPENK

- DLVFSTWDHKAKGHFNCPEGYSGGWWHDECGENNLNGKYNKPRAKSKP
 ERRRGLSWKSQNGRLYSIKSTKMLIHPTDSESFE (SEQ ID NO: 218),
 MFTIKLLLFIVPLVISSRIDQDNSSFDSLSEPKSRF (SEQ ID NO: 219),
 AMLDDVKILANGLLQLGHGLKDFVHKTKGQINDI (SEQ ID NO: 220),
 5 FQKLNIFDQSFYDLSLQTSEIKEEEKELRRTTYKL (SEQ ID NO: 221),
 QVKNEEVKNMSLELNSKLESLLLEEKILLQQKVKYLE (SEQ ID NO: 222),
 EQLTNLIQNQPETPEHPEVTSKTFVEKQDNSIKDL (SEQ ID NO: 223),
 LQTVEDQYKQLNQHSQIKEIENQLRRTSIQEPTE (SEQ ID NO: 224),
 ISLSSKPAPRTTPFLQLNEIRNVKHDGIPAECTT (SEQ ID NO: 225),
 10 IYNRGEHTSGMYAIRPSNSQVFHVYCDVISGSPWTL (SEQ ID NO: 226),
 IQHRIDGSQNFNETWENYKYGFGRLDGEFWLGLEKI (SEQ ID NO: 227),
 YSIVKQSNYVLRIELEDWKDNKHYIEYSFYLGHE (SEQ ID NO: 228),
 TNYTLHLVAITGNVPNAIPENKDLVFSTWDHKAKG (SEQ ID NO: 229),
 HFNCPEGYSGGWWHDECGENNLNGKYNKPRAKSKP (SEQ ID NO: 230),
 15 and/or ERRRGLSWKSQNGRLYSIKSTKMLIHPTDSESFE (SEQ ID NO: 231).

Also preferred are the polynucleotides encoding these polypeptides. The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

- 20 This gene is expressed primarily in liver.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, angiogenesis and neovascularisation associated with tumour
 25 development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., vascular, liver,
 30 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 116 as residues: Arg-18 to Asp-27, Leu-29 to Arg-36, Ser-90 to Tyr-104, Val-108 to
5 Lys-114.

The tissue distribution primarily in liver and homology to angiopoietin-2 indicates that the protein products of this gene are useful for the treatment and/or detection of disorders associated with angiogenesis including the inhibition of angiogenesis and neovascularisation associated with tumour development; the
10 promotion of neovascularisation and wound healing; the treatment of ischaemia; thromboembolytic disease; atherosclerosis; inflammation; and diabetes. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their
15 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
20 related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
25 formula of a-b, where a is any integer between 1 to 1622 of SEQ ID NO:16, b is an integer of 15 to 1636, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

30 Preferred polypeptides of the invention comprise the following amino acid sequence: LPPRGPATFGSPGCPPANSPPSAPATPE PARAPERV (SEQ ID NO: 232). Polynucleotides encoding these polypeptides are also provided.

When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates fibroblast cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. The translation product of this gene shares sequence homology with murine claudin-1 and other murine and human members of the claudin family of integral membrane proteins which are structurally similar and contain four transmembrane domains (e.g., See Genbank Acc. Nos. gil3335182 (AF072127) and/or gil4128015|gnllPIDle1363658). Three integral membrane proteins, claudin-1, -2, and occludin, are known to be components of tight junction (TJ) strands. FLAG-tagged claudin-1 and -2 protein have been demonstrated using immunofluorescence microscopy to be highly concentrated at cell contact sites as planes through a homophilic interaction. It is believed that claudin-1 and -2 are mainly responsible for TJ strand formation, and occludin is an accessory protein in some function of TJ strands (See, J. Cell Biol 143:391-401 (1998), which is hereby incorporated by reference herein).

This gene is expressed primarily in wound healing tissues, and various carcinoma tissues, and, to a lesser extent, in some other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumorigenesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of wounded tissues, and cancerous tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in healing wound tissue and various carcinomas indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of wounds and tumors. Representative uses are described elsewhere herein. Additionally, the homology of the translation product of this gene to claudin-1, a integral membrane protein involved in tight junction formation, and the biological activity of supernatants from cells expressing this gene on fibroblast cells in EGR assays indicate that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of cancer and other proliferative disorders. Expression within cellular sources marked by proliferating cells (e.g., healing wound and various carcinomas) and the homology of the translation product of this gene to a family of claudin proteins suggests that this protein may play a role in the regulation of cellular division and tight junction formation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1242 of SEQ ID NO:17, b is an integer of 15 to 1256, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

The translation product of this gene shares sequence homology with fibulin which is thought to be important in cellular adhesion and extracellular matrix organization.

Preferred polypeptides of the invention comprise the following amino acid sequence: GTRAGVSKYTGGRGVTWAPSSAAVPRISSATMRMGLTSFSTTGA (SEQ ID NO: 233),

WQSGHRLWQLEWPPPPLSADEHPWEGPLPGTSPSPKFSMPSPVPHGHRPTL

5 TMTRSWRIFNNIA

YRSSANRLFRVIRREHGDPLIEELNPGDALEPEGRGTGGVVTDGDLGMLDL
ILSHGESMAQPLSVFRG

NQGFNNWLRVVPRTFRGAFARGAKVVLYTKKSGAHLRIIDGGSGYLCEME
PVAHFGLGKDEASSVEVTW

10 PDGKMVSRNVASGEMNSVLEILYPRDEDTLQDPAPLECGQGFSSQENGHCMDTNECIQFPFVCPRDKPVC VNTYGSYRCRTNKKCSXGLRVPTRMAHTGL (SEQ ID NO: 234), WQSGHRLWQLEWPPPPLSADEHPWEGPLPGTSPSPK (SEQ ID NO: 235), FSPSPVPHGHRPTLTMTRSWRIFNNIAYRSS (SEQ ID NO: 236), ANRLFRVIRREHGDPLIEELNPGDALEPEGRGTGGVV (SEQ ID NO: 237),

15 TDFDGDGMLDLILSHGESMAQPLSVFRGNQGFNN (SEQ ID NO: 238), NWLRVVPRTFRGAFARGAKVVLYTKKSGAHLRIID (SEQ ID NO: 239), GGSGLCEMEPVAHFGLGKDEASSVEVTWPDGKMVS (SEQ ID NO: 240), RNVASGEMNSVLEILYPRDEDTLQDPAPLECGQGF (SEQ ID NO: 241), SQENGHCMDTNECIQFPFVCPRDKPVCVNTYGSYR (SEQ ID NO: 242),

20 and/or CRTNKKCSXGLRVPTRMAHTGL (SEQ ID NO: 243). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 10. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 10.

25 This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, thrombosis, atherosclerosis, neoplasia, schizophrenia, Alzheimer's
30 disease, Parkinson's disease, Huntington's disease, transmissible spongiform encephalopathies (TSE), Creutzfeldt-Jakob disease (CJD), specific brain tumors, aphasia, mania, depression and dementia. Similarly, polypeptides and antibodies

directed to these polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and cardiovascular systems, expression of this gene at significantly higher or lower levels may be

5 routinely detected in certain tissues or cell types (e.g., brain, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or cerebrospinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution in brain and the homology to fibulin suggests that the protein product of this gene is useful for the treatment and diagnosis of developmental, degenerative and/or neoplastic conditions (such as cancer) with mechanisms contingent on the regulation of cellular adhesion and extracellular matrix organization. Fibulin itself, can be used to manipulate adhesion of cells to fibronectin,

15 collagen, laminin, and possibly also other proteins. Thrombosis, atherosclerosis and restenosis may be potential cardiovascular targets for application. In addition polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.

20 Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating

25 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

30 expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or

receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1129 of SEQ ID NO:18, b is an integer of 15 to 1143, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

- The translation product of this gene shares sequence homology with carbonic anhydrase VI, which is thought to be important in protein degradation and pH regulation (see e.g., EMBL locus BTCARANVI, accession X96503; and Jiang et al.,
20 Biochem. J. 318:291-296 (1996) which is hereby incorporated herein, by reference). Based on this homology, it is likely that this gene would have activity similar to carbonic anhydrase.

- Preferred polypeptides of the invention comprise the following amino acid sequence: GQHWTYEGPHGQDHWP (SEQ ID NO: 248), QSPIDIQTDSVTFD
25 (SEQ ID NO: 249), LHNNGHTVQLSLPST (SEQ ID NO: 250), KYVAAQLHLHWG (SEQ ID NO: 251), and/or AELHIVHYDSDSY (SEQ ID NO: 252). Polynucleotides encoding these polypeptides are also provided.

- The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in
30 linkage analysis for chromosome 1.

This gene is expressed primarily in fetal tissues and brain tissue, and, to a lesser extent, in melanocytes, wilms tumor and retinal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, glaucoma and alkalosis resulting from disease of the kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the systems regulating ionic balance and pH in the fluids of the body, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., metabolic, regulatory, renal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 119 as residues: Tyr-24 to His-32, Pro-38 to Ala-44, Pro-66 to Glu-75, His-111 to Gly-116, Tyr-139 to Ser-146, Thr-176 to Ser-181, Lys-239 to Lys-249.

The tissue distribution and homology to secreted carbonic anhydrase suggests that polynucleotides and polypeptides corresponding to this gene are useful for developing drugs that modulate ionic balance in the serum and in the retina, and may be used for treating diseases such as glaucoma or alkalosis secondary to renal disease. Representative uses are described elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1523 of SEQ ID NO:19, b is an integer of 15 to 1537, where both a and b correspond to the positions of nucleotide
5 residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

The translation product of this gene shares sequence homology with murine CD63/ME491 which is thought to be important in activation of macrophage and
10 platelet population (marker of); CD37 (Genbank Acc. No. gil29794), a human leukocyte marker; and several members of the tetraspanin protein family (See, e.g., Genbank Acc. No. gil3152703 (AF065389) and gil2995865 (AF053455)), which are expressed in a wide variety of species and regulate cell adhesion, migration, proliferation and differentiation.

15 The transmembrane 4 superfamily (TM4SF) which has at least 16 members is the second biggest subfamily among CD antigen superfamilies and activation antigens of T-cells. All TM4SF members contain four putative transmembrane domains, two extracellular loops, and two short cytoplasmic tails. They are variously expressed on
20 immature, early, mature, activated lymphocytes, monocytes, macrophages, granulocytes, platelets, eosinophils, basophils, certain leukemic and lymphoma cells, and a variety of other cells and tissues. CD9 cell surface protein is expressed by both hematopoietic and neural cells, and may play a role in intercellular signaling in the immune and nervous system. CD63 is a 53-Kd lysosomal membrane glycoprotein that
25 has been identified as a platelet activation molecule; it plays an important role in cell adhesion of platelets and endothelial cells.

Increased mRNA for CD63 antigen was found in atherosclerotic lesions of Watanabe heritable hyperlipidemic rabbits, suggesting a potential role of CD63 in progression of atherosclerosis. CD63 is also a mast cell marker. This gene also shares close homology with C33 antigen (CD82); CD82 was originally identified as the
30 target of several mAbs inhibitory to syncytium formation induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be a target for the development of a drug for this leukemia.

CD81 is the target of an antiproliferative antibody. A diverse group of human cell lines, including hematolymphoid, neuroectodermal, and mesenchymal cells, express the CD81 protein.

Many of the lymphoid cell lines, in particular those derived from large cell lymphomas, were susceptible to the antiproliferative effects of the antibody. CD81 may therefore play an important role in the regulation of lymphoma cell growth. CD9, CD20, CD37, CD63, CD81 and CD82 have been implicated in the regulation of cell growth, adhesion, and signal transduction of B, T lymphocytes and some other non-lymphoid cells. They associate with CD2, CD21, CD4, CD8, MHC Class II molecules, integrins, and function as co-receptor for T, B and other lymphoid cells. Some TM4SF are leukocyte antigens, highly expressed in activated leukocytes, lymphocytes, and are highly specific surface markers for lymphoblastic leukemia, lymphoma, melanoma, and neuroblastoma. CD9 has been shown to be involved in cell motility and tumor metastasis. These antigens could be a valuable immunogen or target to implement active and passive immunotherapy in patients with cancer. Others have been shown to be involved in inhibition of prostate cancer metastasis.

Preferred polynucleotides of the invention comprise the following nucleic acid sequence:

```
GGCCGCGCCGCGCTGCCGCCGCGCGCGATTCTGCTTCTCAGAAGAT
GCACTATTATAGATACTCTAACGCCAAGGTCAGCTGCTGGTACAAGTACC
TCCTTTTCAGCTACAACATCATCTTCTGATTGGCTGGAGTTGTCTTCCTTGG
AGTCGGGCTGTGGGCATGGAGCGAAAAGGGTGTGCTGTCCGACCTCACCA
AAGTGACCCGGATGCATGGAATCGACCCTGTGGTGCTGGTCCTGATGGTG
GGCGTGGTGATGTTACCCCTGGGGTTCGCCGGCTGCGTGGGGGCTCTGCG
GGAGAATATCTGCTTGCTCAACTTTTTCTGTGGCACCATCGTGCTCATCTT
CTTCCTGGAGCTGGCTGTGGCCGTGCTGGCCTTCCTGTTCCAGGACTGGGT
GAGGGACCGGTTCCGGGAGTTCTTCGAGAGCAACATCAAGTCCTACCGGG
ACGATATCGATCTGCAAAACCTCATCGACTCCCTTCAGAAAGCTAACCAG
TGCTGTGGCGCATATGGCCCTGAAAGACTGGGACCTCAGACGTCTACTTC
AATTGCAGCGGTGCCAGCTACAGCCGAGAGAATGCGGGGTCCCCTTCTCC
TGCTGCGTGCCAGATCCTGCGCAAAAAGTTGTGAACACACAGTGTGGATA
TGATGTCAGGATTCAGCTGAAGAGCAAGTGGGATGAGTCCATCTTCACGA
```

AAGGCTGCATCCAGGCGCTGGAAAGCTGGCTCCCGCGGAACATTTACATT
GTGGCTGGCGTCTTCATCGCCATCTCGCTGTTGCAGATATTTGGCATCTTC
CTGGCAAGGACGCTGATCTCAGACATCGAGGCAGTGAAGGCCGGCCATCA
CTTCTGAGGAGCAGAGTTGAGGGAGCCGAGCTGAGCCACGCTGGGAGGC
5 CAGAGCCTTTCTCTGCCATCAGCCCTACGTCCAGAGGGAGAGGAGCCGAC
ACCCCCAGAGCCAGTGCCCCATCTTAAGCATCAGCGTGACGTGACCTCTC
TGTTTCTGCTTGCTGGTGCTGAAGACCAAGGGTCCCCCTTGTTACCTGCCC
AAACTTGTGACTGCATCCCTCTGGAGTCTACCCAGAGACAGAGAATGTGT
CTTTATGTGGGAGTGGTGACTCTGAAAGACAGAGAGGGCTCCTGTGGCTG
10 CCAGGAGGGCTTGACTCAGACCCCTGCAGCTCAAGCATGTCTGCAGGAC
ACCTGGTCCCCCTCTCCAGTGGCATCCCAAACATCTGCTTTGGGTCCATC
CCACATCTGTGGGTGGGCGGCTGGGTAAGAAGGGAACCCACAGGCGTG
GAACAGGGCATCCTCTCTCCCATCCAAGCAAAGCCAGCATGGGGGCCTGC
CCGTAACGGGAGGCGGACGTGGCCCCGCTGGGCCTCTGAGTGCCAGCGCA
15 GTCTGCTGGGACATGCACATATCAGGGGTTGTTTGCAGGATCCTCAGCCA
TGTTCAAGTGAAGTAAGCCTGAGCCAGTGCGTGGAAGTGGTGCCACGGGAG
TGCCTTGCCACTGTCCCCCTGTGTCCACCAGCTATTCTCCTGGCGCCGGA
ACTGCCTCTGGTCTTGATAGCATTAAAGCCCTGATTGGCCGGTGGCGCGGTG
GGCATGGTTCTTCACTGAGAGCCGGCTCTCCTTTTCTTAAAGTGTGTAAAT
20 AGTTTATTT (SEQ ID NO: 253).

Preferred polypeptides of the invention comprise the following amino acid sequence:

MHYRYRNAKVSCWYKYLFSYNIIFWLAGVVFLGVGLWAWSEKGVLSDL
TKVTRMHGIDPVVLVLMVGVMFTLGFAGCVGALRENICLLNFFCGTIVLIFF
25 LELAVAVLAFLFQDWVRDRFREFFESNIKS YRDDIDLQNLIDSLQKANQCCGA
YGPEDWDLNVYFNCSGASYSREKCGVPFSCCVDPAPQKVNTQCGYDVRIQ
LKSKWDESIFTKGCIQALESWLPRNIYIVAGVFIAISLLQIFGIFLARTLISDIEAV
KAGHHF (SEQ ID NO: 254). Polynucleotides encoding these polypeptides are also
provided.

30 The gene encoding the disclosed cDNA is believed to reside on chromosome
10. Accordingly, polynucleotides related to this invention are useful as a marker in
linkage analysis for chromosome 10.

This gene is expressed primarily in infant and human brain and, to a lesser extent, in pancreas islet cell tumor, Wilm's tumor, uterine cancer, and B cell lymphomas.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: cancers and central nervous system disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
10 particularly of the, immune, metabolic and central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression
15 level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 120 as residues: Met-1 to Ala-9.

The tissue distribution in infant and human brain, and various tumors, and homology to murine CD63/ME491, human CD37, and tetraspanins indicates that the
20 protein product of this gene is useful for the study, detection, treatment, and/or prevention of central nervous system diseases and cancers. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells, and its homology indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of
25 developmental diseases and disorders, cancer, and other proliferative conditions.

Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as
30 occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of

potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, 5 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

The protein is useful in modulating the immune response to aberrant 10 polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional 15 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of 20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2658 of SEQ ID NO:20, b is an 25 integer of 15 to 2672, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

The translation product of this gene shares sequence homology to several 30 steroid receptor proteins (e.g., See Genbank Acc. Nos. gnllPIDle314174, gnllPIDle1154367 (AJ002030), and/or gnllPIDle257707).

Preferred polypeptides of the invention comprise the following amino acid sequence:

SGNLGSADGWAYIDVEVRRPWAFVGP GCSRSSGNGSTAYGLVGSPRWLSPF
HTGGAVSLPRRPRGP

- 5 GPVLGVARPCLRCVLRPEHYEPGSHYSGFAGRDASRAFTGD CSEAGLVDD
VSDLSAAEMLTLHNWLSFY
EKNYVCVGRVTGRFYGEDGLPTALTQVEAAITRGLEANKLQLQEKQTFPPC
NAEWSSARGSRLWCSQKS
GGVSRDWIGVPRKLYKPGAKEPRCVCVRTTGPPSGQMPDNPPHRNRGDLDH
10 PNLAEYTGCPPLAITCSFPL (SEQ ID NO: 255),
SGNLGSADGWAYIDVEVRRPWAFVGP GCSRSSGNGS (SEQ ID NO: 256),
TAYGLVGSPRWLSPFHTGGAVSLPRRPRGPGPVLGV (SEQ ID NO: 257),
ARPCLRCVLRPEHYEPGSHYSGFAGRDASRAFTGD (SEQ ID NO: 258),
CSEAGLVDDVSDLSAAEMLTLHNWLSFYEKNYVCVG (SEQ ID NO: 259),
15 RVTGRFYGEDGLPTALTQVEAAITRGLEANKLQLQ (SEQ ID NO: 260),
EKQTFPPCNAEWSSARGSRLWCSQKSGGVSRDWIGV (SEQ ID NO: 261),
PRKLYKPGAKEPRCVCVRTTGPPSGQMPD (SEQ ID NO: 262), and/or
NPPHRNRGDLDPNLAEYTGCPPLAITCSFPL (SEQ ID NO: 263).

Polynucleotides encoding these polypeptides are also provided.

- 20 This gene is expressed primarily in brain and, to a lesser extent, in variety of other tissues and cell types.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, degenerative and behavioral diseases of the brain such as schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, transmissible spongiform encephalopathies (TSE), Creutzfeldt-Jakob disease (CJD), specific brain tumors, aphasia, mania, depression, dementia, paranoia, addictive behavior and sleep disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
25 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher
30

or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 121 as residues: Glu-42 to Pro-53, Ser-67 to Thr-73, Ala-84 to Leu-90.

The tissue distribution in brain and the homology to steroid receptor proteins indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD), aphasia, specific brain tumors, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1494 of SEQ ID NO:21, b is an integer of 15 to 1508, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in kidney and gall bladder tissues, fetal tissue, and testes tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal disorders, metabolic diseases, and disorders of the reproductive and developing organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal, metabolic, developing, and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., renal, metabolic, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 122 as residues: Lys-60 to Ala-66.

The tissue distribution in kidney and gall bladder tissues, testicular tissue, and fetal tissues, suggests that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the renal system, reproductive system, metabolic system and developing systems. Furthermore, the tissue distribution in kidney suggests that this gene or gene product is useful in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome.

Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g., endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1433 of SEQ ID NO:22, b is an integer of 15 to 1447, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 13

Preferred polypeptides of the invention comprise the following amino acid sequence: RDNDYLLHGHRRPPMF (SEQ ID NO: 264),
SFRACFKSIFRIHTETGNIWTHLL (SEQ ID NO: 265), and/or
GFVLFLFLGILTMLRPNMYFMAPLQEKVV (SEQ ID NO: 266). Polynucleotides
10 encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in bone marrow, fetal liver and spleen
15 tissues, several types of leukocytes including neutrophils, and T-cells, placental tissue, and brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, diseases and/or disorders of the immune system and central nervous system including AIDS, Lupus, hemotological cancers, mood disorders, and dementia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of
25 the immune system and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression
30 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 123 as residues: Glu-24 to Tyr-35, Arg-83 to Thr-92, Pro-148 to Gly-154.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in fetal liver and spleen tissues, and several types of leukocytes, suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or

receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1569 of SEQ ID NO:23, b is an integer of 15 to 1583, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

The translation product of this gene shares sequence homology with gp25L, which is thought to be important in protein processing.

This gene is expressed primarily in stimulated synovium, cerebellum, and
20 placental tissues, and, to a lesser extent, in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, disorders of developing systems, central nervous system,
25 and musculo-skeletal system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, central nervous system, musculo-skeletal, and developing systems, expression of this gene at significantly higher or lower levels
30 may be routinely detected in certain tissues or cell types (e.g., immune, neural, musculo-skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample

taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to gp25L suggests that the protein
5 product of this gene is useful for treatment and/or diagnosis of disorders of immune, central nervous system, musculo-skeletal, and developing systems. In addition, the expression of this gene product in synovium suggests a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g., arthritis, trauma,
10 tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e., spondyloepiphyseal dysplasia congenita, familial arthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia
15 type Schmid). Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1655 of SEQ ID NO:24, b is an integer of 15 to 1669, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in immune and hematopoietic cells, and breast and brain tissues, and, to a lesser extent, in several other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic disorders, disorders of the central nervous system and reproductive organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoietic, central nervous system and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, reproductive, neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in breast, brain, and immune tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the immune, hematopoietic, central nervous and reproductive systems. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1039 of SEQ ID NO:25, b is an integer of 15 to 1053, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

Preferred polypeptides fragments from this alternative reading frame comprise:

- 10 TGPEFPGSNSTVARRIKDLAADIEEELVCRLKICDGFSLQLDESADVSGLA VLL
VFVRYRFNKSIEED
LLLCESLQSNATGEEIFNCINSFMQKHEIEWEKCVDVCSDASRAVDGKIAEAV
TLIKYVAPESTSSHCLL
YRHALAVKIMPTSLKNVLDQAVQIINYIKARPHQSRLKILCEEMGAQHTALL
- 15 LNTEVRWLSRGKVLVRL
FELRRELLVFMDSAFRLSDCLTNSSWLLRLAYLADIFTKLNEVNLSMQGKNV
TVFTVFDKMSSLLRKLEF
WASSVEEENFDCFPTLSDFLTEINSTVDKDICAIVQHLRGLRATLLKYFPVTN
DNNAWVRNPFTVTVKP
- 20 ASLVARDYESLIDLTSDSQVKQNFSELSLNDFWSSLIQEYPSIARRAVRVLLPF
ATMHL CETGFSYYAAT
KTKYRKRLDAAPHMRIRLSNITPNIKRICDKKTQKHCSH (SEQ ID NO: 267),
DIEEELVCRLKICDGFSLQLDESADVSGLA V (SEQ ID NO: 268),
NSFMQKHEIEWEKCVDVCSDASRAVDGKIAEAVTLI (SEQ ID NO: 269),
- 25 LDQAVQIINYIKARPHQSRLKILCEEMGAQHTALL (SEQ ID NO: 270),
SAFRLSDCLTNSSWLLRLAYLADIFTKLNEVNLSMQGKNVTVFTVFDKM
(SEQ ID NO: 271), SDFLTEINSTVDKDICAIVQHLRGLRATLLK (SEQ ID NO:
272), and/or SDSQVKQNFSELSLNDFWSSLIQEYPSIARRAVRVLLP (SEQ ID
NO: 273). Also preferred are polynucleotide fragments encoding these polypeptide
- 30 fragments.

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in spleen from a chronic lymphocytic leukemia patient, and hodgkin's lymphoma, and, to a lesser extent, in pancreatic islet cell tumors and activated T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia; hodgkin's lymphoma; pancreatic islet cell cancer; cancer in general; hematopoietic disorders; immune dysfunction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and pancreas, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spleen from a chronic lymphocytic leukemia patient, and hodgkin's lymphoma, pancreatic islet cell tumors, and activated T-cells. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the protein product of this gene is useful for the diagnosis and/or treatment of a variety of cancers, including CLL; Hodgkin's lymphoma; and pancreatic cancer. Expression of this gene product in a variety of cancers suggests that it may be a bad player and may likely be a target for inhibitors as therapeutics. Alternately, this gene product may be expressed in both normal and abnormal hematopoietic tissues, where it may play necessary roles in the proliferation; survival; differentiation; or activation of hematopoietic cell lineages. Likewise, expression in pancreatic islet cell tumors

may simply reflect a necessary role that this protein plays in normal pancreatic function. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
5 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of
10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1463 of SEQ ID NO:26, b is an
15 integer of 15 to 1477, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

When tested against U937 Myeloid cell lines, supernatants removed from cells
20 containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent other cells, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the
25 differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in endometrial tumor tissue and cartilage tissue, and to a lesser extent in several other tissues and organs.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to tumors and disorders of the musculo-skeletal system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the musculo-skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., musculo-skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 127 as residues: Met-1 to Ser-8.

The tissue distribution in musculo-skeletal tissues and biological activity in the GAS assay, suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the musculo-skeletal system, and cancers thereof. In addition, the expression of this gene product in cartilage tissue suggests a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g., arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e., spondyloepiphyseal dysplasia congenita, familial arthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
5 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2490 of SEQ ID NO:27, b is an integer of 15 to 2504, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 18

The gene encoding the disclosed cDNA is thought to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in breast and cerebellum tissues, as well as in
15 cells of the hematopoietic system, and, to a lesser extent, in several other organs and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, disorders of the brain, reproductive system and hematopoietic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic system, central nervous system and reproductive system,
25 expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, neural, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression
30 level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 128 as residues: Gly-56 to Gly-86, Leu-107 to Ala-112, Ala-121 to Thr-129, Lys-164 to Gln-174.

5 The tissue distribution in immune, reproductive, and neural tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the immune and haemopoietic system, the central nervous system, and the reproductive system. Furthermore, the expression in the breast tissue may indicate its uses in breast neoplasia and breast cancers, such as fibroadenoma, papillary carcinoma, ductal carcinoma, Paget's disease, medullary
10 carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases.

Alternatively, the tissue distribution in cerebellum tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the
15 detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the
20 gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. In addition, the tissue distribution in immune system cells and tissues suggests that the translation product of this gene is useful for the detection and/or treatment of immune system disorders. This gene product may be involved in the
25 regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker
30 and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory

bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1852 of SEQ ID NO:28, b is an integer of 15 to 1866, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

The translation product of this gene shares weak sequence homology with dehydrogenase enzymes (See, e.g., gnlIPIDle1316908) which are thought to be important in a variety of enzymatic conversions, including the biosynthesis of clavulanic acid from a precursor clavulanic acid aldehyde. The obtained clavulanic acid is in turn a key ingredient in antibiotics.

Preferred polypeptides of the invention comprise the following amino acid sequence: DSRISLLVNNAGVGATASLLESDADK (SEQ ID NO: 274). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in CD34 positive hematopoietic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic diseases and/or disorders; impaired immune function; lymphomas & leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:

129 as residues: Pro-97 to Pro-113.

The tissue distribution in CD34 positive hematopoietic cells indicates that the protein product of this gene is useful for the diagnosis and/or treatment of a variety of hematopoietic disorders. Expression of this gene product specifically in CD34 positive cells suggests that it plays a role in early events of hematopoiesis, including proliferation; survival; differentiation; and activation of early stem and committed progenitor cells. The protein product of this gene is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1487 of SEQ ID NO:29, b is an integer of 15 to 1501, where both a and b correspond to the positions of nucleotide

10 residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

Preferred polypeptides of the invention comprise the following amino acid sequence:

15 GTPAGTGPEFFGRPTRPSRTESAQTTQHSPLRPLWRLKRDSSPCHPQTRADWG
VCPWGGAAQGLRPGCH
LAPRRCLCPGSCCPWHWAEAQWSFLWRGLWGLRTLPTALRASPAASGTVTY
SACLGTSCLLRAPCWRLRT CRQSWC (SEQ ID NO: 275),
GTPAGTGPEFFGRPTRPSRTESAQTTQH (SEQ ID NO: 276),
20 SPLRPLWRLKRDSSPCHPQTRADWGVCPW (SEQ ID NO: 277),
GGAAQGLRPGCHLAPRRCLCPGSCCPWHWA (SEQ ID NO: 278),
EAQWSFLWRGLWGLRTLPTALRASPAASGT (SEQ ID NO: 279), and/or
VTYSACLGTSCLLRAPCWRLRTCRQSWC (SEQ ID NO: 280). Polynucleotides
encoding these polypeptides are also provided.

25 The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in osteoarthritis, breast cancer, and uterine cancer, and, to a lesser extent, in brain.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, cancer, particularly breast and uterine cancer; and neurological diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above

5 tissues or cells, particularly of the breast, lymph node, and CNS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, breast, skeletal, joint, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an

10 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 130 as residues: Gln-75 to Cys-80.

15 The tissue distribution in breast and uterine cancer indicates that the protein product of this gene is useful for the diagnosis and/or treatment of a variety of cancers, particularly breast cancer and uterine cancer. Expression of this gene in brain also indicates that it may play a role in neurological function, and that its absence may lead to disorders such as Alzheimer's & Parkinson's disease. Expression of this gene

20 product at elevated levels within cancerous tissue indicates that it may be a player in the progression of the disease, perhaps by driving proliferation or blocking differentiation or apoptosis. Therefore, beneficial therapeutics may be developed based upon attempts to block this gene product.

Representative uses are described in the "Hyperproliferative Disorders" and

25 "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain

30 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may

also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue
5 differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the
10 protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
20 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1738 of SEQ ID NO:30, b is an integer of 15 to 1752, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene shares sequence homology with a yeast hypothetical 52.9 KD protein CDC26-YMR31 intergenic region (See Genbank Accession No. gplD50617/YSCCHRVI_114.). This gene has been mapped to chromosome 18q22-
30 23, and therefore can be used in linkage analysis as a marker for 18q22-23.

This gene is expressed primarily in whole brain tissue, as well as brain specific tissues such as hypothalamus, frontal cortex, cerebellum, amygdala, and hippocampus tissues, as well as other brain specific tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, schizophrenia, developmental disorders, and abnormal mental states. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, brain, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
15 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
131 as residues: Met-98 to Gln-107, Gly-120 to Gly-126, Pro-138 to Trp-145, Leu-
159 to Gly-169, Val-211 to Arg-217, Cys-256 to His-262, Glu-320 to Val-327, Phe-
20 399 to Asn-406, Asp-444 to Ser-450, Asp-475 to Trp-488.

The tissue distribution in whole brain tissue and brain specific tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for treating and/or diagnosing neural and neurodegenerative disorders. Furthermore, the tissue distribution suggests that polynucleotides and polypeptides corresponding to
25 this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance,
30 and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Elevated expression of this gene product within

the frontal cortex of the brain suggests that it may be involved in neuronal survival; synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. Additionally, the amygdala processes sensory information and relays this to other areas of the brain including the endocrine and autonomic domains of the hypothalamus and the brain stem. Thus, the translation product of this gene may also be useful for the detection and/or treatment of neural disorders that impact processes mediated by the amygdala. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2138 of SEQ ID NO:31, b is an integer of 15 to 2152, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

Preferred polypeptides of the invention comprise the following amino acid sequence: PPRPSTSGQWG (SEQ ID NO: 281) and/or RRPFTSAQTG (SEQ ID NO: 282). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in breast and soleus tissues, and, to a lesser extent, in several cell types, including T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, breast cancer, and musculo-skeletal diseases and/or disorders.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lactation system and breast, as well as the musculo-skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., musculo-skeletal, breast, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 132 as residues: Thr-35 to Lys-43, Pro-59 to Arg-64.

The tissue distribution in soleus tissue indicates that the protein product of this gene is useful for the detection, treatment, and/or prevention of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stroke, angina, thrombosis, and wound healing. Representative uses are described elsewhere herein. Alternatively, the expression in the breast tissue may indicate its uses in breast neoplasia and breast cancers, such as fibroadenoma, papillary carcinoma, ductal carcinoma, Paget's disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1743 of SEQ ID NO:32, b is an
 5 integer of 15 to 1757, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

The gene encoding the disclosed cDNA is believed to reside on chromosome
 10 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

Preferred polypeptides of the invention comprise the following amino acid sequence: GTGWDFGLAAVCLRAAEVAGSFK (SEQ ID NO: 283),
 GYRRVFEEYMRVISQRYPDRIEGENYLPQPIYRHIASFLSVFKLVLIIGLIIVGK
 15 DPFAFFGMQAPSI
 WQWGQENKVYACMMVFFLSNMIENQCMSTGAFEITLNDVPVWSKLESGHL
 PSMQQLVQILDNEMKLVHMDSIPHRS (SEQ ID NO: 284),
 GYRRVFEEYMRVISQRYPDRIEGENYLPQPIYR (SEQ ID NO: 285),
 HIASFLSVFKLVLIIGLIIVGKDPFAFFGMQAPSI (SEQ ID NO: 286),
 20 WQWGQENKVYACMMVFFLSNMIENQCMSTGAFEI (SEQ ID NO: 287),
 TLNDVPVWSKLESGHLPSMQQLVQILDNEMKLVHMDSIPHRS (SEQ ID NO: 288),
 and/or DSIPHRS (SEQ ID NO: 289). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fast-growing tissues such as early
 25 development stage tissues, cancerous tissues, and hematopoietic tissues, and, to a lesser extent, in some other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
 30 not limited to, growth disorders, tumorigenesis, and immune and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s)

or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fast-growing tissues such as early development stage tissues, cancer tissues, and hematopoietic tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded
5 tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fast-growing tissues such as early development stage
10 tissues, cancerous tissues, and hematopoietic tissues, indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of growth disorders, tumorigenesis, and immune and inflammatory disorders. Similarly, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of cancer and other
15 proliferative disorders. Expression in cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues suggests that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene may be useful in the treatment
20 of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1452 of SEQ ID NO:33, b is an

integer of 15 to 1466, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

5 Preferred polypeptides of the invention comprise the following amino acid sequence: GRARGRPPGPEAAPASLSVSLRREVHSRGE (SEQ ID NO: 290).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in olfactory epithelium.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, olfactory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
15 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the olfactory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., olfactory, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
20 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 134 as residues: His-24 to Ala-29, Glu-42 to Glu-49.

The tissue distribution primarily in the olfactory epithelium indicates a role for
25 this protein in the treatment and/or diagnosis of olfactory and sensory disorders, including loss of the sense of smell. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 512 of SEQ ID NO:34, b is an integer of 15 to 526, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

This gene is expressed primarily in 8 week embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly during fetal development, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., embryonic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The expression of this gene primarily in the embryo, indicates a key role for this protein in embryo development and further indicates its usefulness in the treatment and/or detection of embryonic developmental defects.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2398 of SEQ ID NO:35, b is an integer of 15 to 2412, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 26

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 136 as residues: Trp-25 to Thr-38, Pro-83 to Ala-88.

The tissue distribution in neutrophils suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of immune system disorders, especially those affecting neutrophils. Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory

bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a
5 tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1260 of SEQ ID NO:36, b is an integer of 15 to 1274, where both a and b correspond to the positions of nucleotide
15 residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in
20 linkage analysis for chromosome 1.

This gene is expressed primarily in fetal liver and brain tissues, and, to a lesser extent, in various other fetal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, developmental and neurological conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
30 developing, immune and central nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, developing, neural, cancerous and wounded tissues) or bodily

fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 137 as residues: Pro-75 to Asn-81, Gln-106 to Cys-111, Glu-130 to Asp-141, Arg-176 to Asp-182, Ala-201 to Trp-206, Lys-238 to Thr-246.

The tissue distribution in fetal liver and brain tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, 10 detection and/or treatment of growth disorders and neoplasias of the immune and central nervous systems. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.

- 15 Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating 20 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated 25 expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the 30 developing embryo, or sexually-linked disorders.

Alternatively, expression of this gene product in fetal liver/spleen tissue suggests a role in the regulation of the proliferation; survival; differentiation; and/or

activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 5 Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory
- 10 bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify
- 15 agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
- 20 related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
- 25 formula of a-b, where a is any integer between 1 to 1022 of SEQ ID NO:37, b is an integer of 15 to 1036, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

- 30 This gene shares sequence homology to fibulin (See GeneSeq Accession No. R11148 and R11149). Fibulin binds to the cytoplasmic domain of the beta-1 subunit of integrin adhesion receptors in a cation-dependent, EDTA-reversible manner. Thus,

this gene may be used to manipulate adhesion of cells to fibronectin, collagen, laminin, and possibly also other proteins. When tested against both U937 Myeloid cell lines and Jurkat T-cell cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates both T-cells
5 and myeloid cells, and to a lesser extent other tissues and cell types, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT
10 pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

The gene encoding the disclosed cDNA is thought to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

15 This gene is expressed primarily in cerebellum tissue, and, to a lesser extent, in multiple tissues and cell types including prostate, liver, T-cells, kidney, and lung tissues, as well as musculo-skeletal tissues such as endothelial tissue, healing groin wound tissue, fetal heart tissue, and osteosarcoma tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the central nervous system, including dementia, mood disorders, both unipolar and bipolar depression, and Alzheimer's disease, as well as disorders of the musculo-skeletal, renal, and pulmonary systems.
25 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, renal, pulmonary system, and musculo-skeletal system, expression of this gene at significantly higher or lower levels may be routinely
30 detected in certain tissues or cell types (e.g., neural, musculo-skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such

a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 138 as residues: Pro-28 to Thr-45, Arg-59 to Gly-67, Ala-71 to Glu-84, Lys-120 to Asp-126, Pro-159 to Gly-164, Glu-167 to Gly-186, Arg-217 to Ala-225, Glu-245 to Ala-255, Gly-282 to Gly-297, Pro-312 to Gly-324, Thr-356 to Lys-364, Gly-366 to Thr-372, Lys-377 to Ala-383, Gly-397 to Thr-407, Thr-419 to Gly-433.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of cancers, most notably cancers of the central nervous system, pulmonary, and renal systems, as well as the disorders of the central nervous system listed above. Representative uses are described in the "Hyperproliferative Diseases", "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the expression of this gene product in a variety of systems suggests that this gene may be a player in the progression of these diseases, and may be a beneficial target for inhibitors as therapeutics.

Alternatively, the tissue distribution in musculo-skeletal tissues, as the homology to fibulin, suggests that the translation product of this gene is useful for the detection and/or treatment of disorders involving the vasculature. Elevated expression of this gene product by endothelial cells suggests that it may play vital roles in the regulation of endothelial cell function; secretion; proliferation; or angiogenesis. Alternately, this may represent a gene product expressed by the endothelium and transported to distant sites of action on a variety of target organs. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1365 of SEQ ID NO:38, b is an
 5 integer of 15 to 1379, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

The translation product of this gene shares sequence homology with coxsackie
 10 and adenovirus receptor in mouse. Particularly, this gene shares sequence homology with a human A33 antigen, which is a transmembrane protein and a novel member of the immunoglobulin superfamily. (See Proc. Natl. Acad. Sci. U.S.A. 94, 469-474 (1997); see also, Accession No. 1814277; all references available through the accession and reference are hereby incorporated herein by reference.) Therefore, this
 15 gene likely has activity similar to the human A33 antigen.

Preferred polypeptides of the invention comprise the following amino acid sequence:

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAW
 YTLHGEVSSSQPWEVPMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMP
 20 SRNLSLRLEGLQEKDSGPYSCSVNVQNKQGKSRGHSIKTLELNVLVPPAPPSC
 RLQGVPHVGANVTLSQSPRSKPAVQYQWDRQLPSFQTFAPALDVIRGSL
 LTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAADVAGAVVGTLVG
 LGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTLSSV
 TSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQIPISPIPGGVSSSG
 25 LSRMGAVPVMVPAQSQAGSL (SEQ ID NO: 291),
 MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHL (SEQ ID NO: 292),
 PANRLQAVEGGEVVLPAWYTLHGEVSSSQPWEVPM (SEQ ID NO: 293),
 VMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMP (SEQ ID NO: 294),
 SRNLSLRLEGLQEKDSGPYSCSVNVQNKQGKSRGH (SEQ ID NO: 295),
 30 SIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSQ (SEQ ID NO: 296),
 SPRSKPAVQYQWDRQLPSFQTFAPALDVIRGSL (SEQ ID NO: 297),
 LTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGP (SEQ ID NO: 298),

GAADVAGAVVGTLVGLGLLAGLVLLYHRRGKALEE (SEQ ID NO: 299),
PANDIKEDAIAPRTLWPWKSSDTISKNGTLSSVTS (SEQ ID NO: 300),
ARALRPPHGPPrPGALTPTPSLSSQALPSPRLPTT (SEQ ID NO: 301), and/or
DGAHPQPISIPGGVSSSGLSRMGAVPVMVPAQSQAGSL (SEQ ID NO: 302).

- 5 Polynucleotides encoding these polypeptides are also provided.

This gene is expressed in various tissues including placenta, brain, heart, muscle, adipocytes, and liver.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: viral diseases, and immune diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., CNS, reproductive, vascular, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution in various tissues including placenta, brain, heart, muscle, adipocytes, and liver, and the homology to A33 antigen indicates that the protein product of this gene is useful for the diagnosis and/or treatment of a variety of cancers, most notably cancers of the immune system, as well as viral infections. Expression of this gene product suggests that this gene may be a player in the progression of these diseases, and may be a beneficial target for inhibitors as therapeutics. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:39, b is an integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

Preferred polypeptides of the invention comprise the following amino acid sequence: GSSFVVSEGSYLDISDWLNPAKLSLYY (SEQ ID NO: 303), LDISDWLNPAKL (SEQ ID NO: 304), SDWLNPAKLSL (SEQ ID NO: 305), and/or DACEQLCDPETGE (SEQ ID NO: 310). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in human ovary and adrenal gland tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive diseases and/or disorders, particularly ovarian cancer.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in ovary tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating reproductive system disorders including ovarian cancer, as well as cancers of other tissues where expression has been observed. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1416 of SEQ ID NO:40, b is an integer of 15 to 1430, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

This gene is expressed primarily in thymus and stromal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, aberrant immune responses, such as either chronic or acute inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above

tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample
5 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in thymus stromal cells suggests that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating
10 disorders of the immune system, particularly those involving a pathological inflammatory response. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Furthermore, the gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection,
15 inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify
20 agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
30 formula of a-b, where a is any integer between 1 to 1393 of SEQ ID NO:41, b is an integer of 15 to 1407, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

Preferred polypeptides of the invention comprise the following amino acid sequence: EGKIKICEKKAIVILHTCNS (SEQ ID NO: 311). Polynucleotides
5 encoding these polypeptides are also provided.

This gene is expressed primarily in frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, central nervous system (CNS) diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS, expression of this gene at significantly higher or lower levels may be routinely
15 detected in certain tissues or cell types (e.g., brain, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or cerebrospinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 142 as residues: Pro-41 to Asp-47.

The tissue distribution in frontal cortex indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of CNS disorders including disorders of the brain and nervous system. Representative uses are
25 described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Elevated expression of this gene product within the frontal cortex of the brain suggests that it may be involved in neuronal survival, synapse formation, conductance, neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. It may also
30 be useful in the treatment of such neurodegenerative disorders as schizophrenia, ALS, or Alzheimer's. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors,

to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
10 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 936 of SEQ ID NO:42, b is an integer of 15 to 950, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 33**

This gene is expressed primarily in adipose tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, obesity, Nasu-Hakola disease, cardiovascular disease, non-insulin-dependent diabetes mellitus. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adipose, expression of this gene at significantly
25 higher or lower levels may be routinely detected in certain tissues or cell types (e.g., adipose, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the
30 disorder.

The tissue distribution in adipose suggests that the protein product of this gene is useful for the treatment and diagnosis of metabolic disorders related to lipids and

adipose tissue, such as obesity, Nasu-Hakola disease (membranous lipodystrophy), cardiovascular disease, lipidemia, non-insulin-dependent diabetes mellitus, stroke and carcinoma. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 990 of SEQ ID NO:43, b is an integer of 15 to 1004, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

Preferred polypeptides of the invention comprise the following amino acid sequence: NSARVEFFIPPLRITQKVRSTKS (SEQ ID NO: 312). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is apparently expressed exclusively in IL-1- and LPS-induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, abnormal immune reactions or disorders including, but not limited to, chronic or cyclic neutropenia, neutrophilia, and neutrocytosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression

of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of immune disorders or abnormal reactions mediated by neutrophils, including infection, inflammation, allergy, immunodeficiency, chronic or cyclic neutropenia, neutrophilia, and neutrocytosis, and the like. Moreover, the expression of this gene product suggests a role in regulating the proliferation, survival, differentiation, and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity, immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1667 of SEQ ID NO:44, b is an integer of 15 to 1681, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

The translated ORF of the contig has homology with the human, porcine, and bovine INS10 double-chain insulin precursor, especially around a region containing multiple cysteine residues.

Preferred polypeptides of the invention comprise the following amino acid sequence:
 MMVWNLFPFCFPPLLLLQFIDCQQSSEIEQGFRSLLGHPIFFCPDPCWQSCMN
 CVILSVLSFFFLIRWISKIVAVQKLESSRRKPILFLISCEIASFIHLFLSQMSAEC
 CCFYLVLICKY (SEQ ID NO: 313), MMVWNLFPFCFPPLLLLQFIDCQQSSEIE
 (SEQ ID NO: 314), QGFRSLLGHPIFFCPDPCWQSCMNCVI (SEQ ID NO: 315),
 LSVLSFFFLIRWISKIVAVQKLESSRRKPILFLI (SEQ ID NO: 316), and/or
 ISCEIASFIHLFLSQMSAEC CCFYLVLICKY (SEQ ID NO: 317). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in cells and tissues isolated from a 15 days post-incision healing abdomen wound and, to a lesser extent, in many connective tissues/cells with proliferative capacity, such as osteoclastoma, ovarian cancer, B-cell lymphoma and hepatocellular tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, wound healing, diabetes mellitus, and cancers of the bone and

connective tissues, lymphomas, and cancers of the liver. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly those of the cells and tissues involved in healing tissue damages and regeneration, diabetes mellitus, and many cancers including, but not limited to ovarian cancer, breast cancer, colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, and the like, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 145 as residues: Gln-22 to Phe-31.

The tissue distribution in healing wound and regenerating tissues/cells suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of tissue damages, trauma, necrosis, and tissue regeneration. In addition, since this gene exhibits homology with an insulin precursor, it can be used to regulate the metabolism of glucose or other sugars, the synthesis of proteins, and the formation and storage of neutral lipids.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1351 of SEQ ID NO:45, b is an

integer of 15 to 1365, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

- 5 Preferred polypeptides of the invention comprise the following amino acid sequence:
KVDTPRRHFCPEISFFLTPLPQSARNSTVRNALSGLKNLTPAMISTVSKQDTSK
LGEEE (SEQ ID NO: 318). Polynucleotides encoding these polypeptides are also provided.
- 10 When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large,
- 15 signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.
- This gene is expressed primarily in B-cell lymphoma.
- 20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B-cell lymphoma, immunodeficient or auto-immune conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
- 25 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal
- 30 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution suggests that the protein product of this gene is useful for the detection, treatment, and/or prevention of B-cell lymphomas, as well as other immune disorders including: leukemias, auto-immunities, immunodeficiencies (e.g., AIDS), immuno-suppressive conditions (transplantation) and hematopoietic disorders, such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia, since stromal cells are important in the production of cells of hematopoietic lineages. In addition, this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, the biological activity of supernatants from cells expressing this gene in the GAS assay indicates that this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1123 of SEQ ID NO:46, b is an integer of 15 to 1137, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

The gene encoding the disclosed cDNA is thought to reside on chromosome 10. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 10.

This gene is expressed primarily in infant brain and caudate nucleus tissues, and, to a lesser extent, in various other normal and transformed cell types, including smooth muscle and adult heart tissues, and T-cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and growth defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in infant brain tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, detection and/or treatment of infant and general nervous system disorders and neoplasias. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive

disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2749 of SEQ ID NO:47, b is an integer of 15 to 2763, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of this gene shares weak homology with O-linked GlcNAc transferases (See, e.g., Genbank Acc. No. gil2266994) which are important for a variety of cellular functions, including, secreted protein stability and proper function.

Preferred polypeptides of the invention comprise the following amino acid sequence: LLLCPWWLCFDWS (SEQ ID NO: 319),

MGCIPLIKSISDWRVIALAALWFCLIGLICQALCEDGHKRRILTLGLGFLVIPF
 LPASNLFVRVGFVVAECVLYLPSIGYCVLLTFGFGALSKHTKKKKLIAAVVLG
 ILFINTLRCVLRRTAKWRSEEQLFRSALSVCPLNAKVHYNIGKNLADKGNQTA
 AIRYYREAVRLNPKYVHAMNNLGNILKERNELQEAEEELSLAVQIQPDFAAA
 5 WMNLGIVQNSLKRFEAEQNYRTAIKHRRKYPDCYYNLGRLVRTGCPVPVE
 GKMGYFS (SEQ ID NO: 320),
 MGCIPLIKSISDWRVIALAALWFCLIGLICQALCEDG (SEQ ID NO: 321),
 HKRRILTLGLGFLVIPFLPASNLFFRVGFVVAECVLYL (SEQ ID NO: 322),
 PSIGYCVLLTFGFGALSKHTKKKKLIAAVVLGILFINT (SEQ ID NO: 323),
 10 LRCVLRRTAKWRSEEQLFRSALSVCPLNAKVHYNIGKNL (SEQ ID NO: 325),
 ADKGNQTAAIRYYREAVRLNPKYVHAMNNLGNILKERN (SEQ ID NO: 326),
 ELQEAEEELSLAVQIQPDFAAAWMNLGIVQNSLKRFE (SEQ ID NO: 327),
 and/or AEQNYRTAIKHRRKYPDCYYNLGRLVRTGCPVPVEGKMGYFS (SEQ
 ID NO: 328). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed primarily in substantia nigra and, to a lesser extent, in
 amygdala and brain, striatum.

Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 20 not limited to, neurodegenerative disorders. Similarly, polypeptides and antibodies
 directed to these polypeptides are useful in providing immunological probes for
 differential identification of the tissue(s) or cell type(s). For a number of disorders of
 the above tissues or cells, particularly of the central nervous system and brain,
 expression of this gene at significantly higher or lower levels may be routinely
 25 detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or
 bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or
 another tissue or cell sample taken from an individual having such a disorder, relative
 to the standard gene expression level, i.e., the expression level in healthy tissue or
 bodily fluid from an individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
 148 as residues: Ser-35 to Arg-41.

The tissue distribution in substantia nigra and, to a lesser extent, in amygdala and brain, striatum, suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.

5 Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia,
10 trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

15 In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate
20 ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1562 of SEQ ID NO:48, b is an

integer of 15 to 1576, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

5 This gene is expressed primarily in epithelial-TNF α and INF induced cells and brain frontal cortex.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely
15 detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 149 as residues: Lys-35 to Asp-41, Glu-49 to Leu-63.

 The tissue distribution in the brain suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of neurodegenerative disorders, especially those involving the frontal cortex. Representative uses are
25 described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the elevated expression of this gene product within the frontal cortex of the brain suggests that it may be involved in neuronal survival; synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. It
30 may also be useful in the treatment of such neurodegenerative disorders as schizophrenia; ALS; or Alzheimer's. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate

ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1334 of SEQ ID NO:49, b is an integer of 15 to 1348, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

Preferred polypeptides of the invention comprise the following amino acid sequence: PTRPPTRPLSFTFTKQTSSTCLSLHF (SEQ ID NO: 329).

Polynucleotides encoding these polypeptides are also provided.

- 20 The gene encoding the disclosed cDNA is believed to reside on chromosome 18. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 18.

This gene is expressed primarily in infant brain, frontal cortex, and, to a lesser extent, in melanocytes.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
30 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely

- detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or
- 5 bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 150 as residues: Val-40 to Cys-47, Lys-49 to Gly-54.

- The tissue distribution suggests that the protein product of this gene is useful for the detection, treatment, and/or prevention of neurodegenerative disorders
- 10 especially those involving the frontal cortex. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and
- 15 "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,
- Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia,
- 20 mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

- In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is
- 25 involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the
- 30 protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1250 of SEQ ID NO:50, b is an integer of 15 to 1264, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene shows structural homology with the duck insulin precursor which is thought to be important in metabolic homeostasis (See Accession No. pirlA01600IIPDK insulin precursor). Preferred polypeptide fragments comprise the
15 amino acid sequence:
LECVLLICFRAMSAIYTHTSIGNAQKLFTDGSAFRRVREPLPKEGKSWPQ
(SEQ ID NO: 330). Also preferred are polynucleotide fragments encoding this polypeptide fragment.

This gene is expressed primarily in eosinophil-IL5 induced cells, and, to a
20 lesser extent, in B cell lymphoma, breast lymph node, and CD34 depleted buffy coat (cord blood).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, immune diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain
30 tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal

fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:

5 151 as residues: Arg-39 to Glu-56.

The tissue distribution in hematopoietic tissues suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of immune disorders especially those involving eosinophils and B-cells. The protein product of this gene is useful for the detection, treatment, and/or prevention of a variety of
10 immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene
15 product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an
20 agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host
25 diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene
30 product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological

activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1646 of SEQ ID NO:51, b is an integer of 15 to 1660, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

Preferred polypeptides of the invention comprise the following amino acid sequence: KQNLTNLDVPVQYHVALSDKVK (SEQ ID NO: 331). Polynucleotides encoding these polypeptides are also provided.

20 This gene is expressed primarily in pineal gland and, to a lesser extent, in multiple sclerosis cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, insomnia, multiple sclerosis, and other neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and endocrine system,
30 expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:

5 152 as residues: Pro-7 to Gly-12.

The tissue distribution primarily in pineal gland and, to a lesser extent, in multiple sclerosis cells suggests that the protein product of this gene is useful for treatment of insomnia and jet lag through agonist or antagonist interaction with pineal gland receptors to allow regulation of melatonin production. Representative uses are
10 described elsewhere herein. This gene may also be useful in the treatment of multiple sclerosis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
15 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1664 of SEQ ID NO:52, b is an integer of 15 to 1678, where both a and b correspond to the positions of nucleotide
25 residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

The gene encoding the disclosed cDNA is believed to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in
30 linkage analysis for chromosome 2.

Preferred polypeptides of the invention comprise the following amino acid sequence:

PSCPPPEMKKELPVDSCCLPRSLLEHPQKMDPKRQHIQLLSSLTECLTVDPPLSASV
WRQLYPKHLSQSSLLL

XHLLSSWEQIPKKVQKSLQETIQSLKLTNQELLRKGSSNNQDVVTCD (SEQ ID
NO: 332). Also preferred are the polynucleotides encoding these polypeptides.

5 This gene is expressed primarily in ovary tumors and breast cancer and, to a lesser extent, in normal lung and colon tumors.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, cancer, particularly of the ovary and breast; and colon. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the colon, breast, or female reproductive system, expression of this gene at significantly
15 higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from
20 an individual not having the disorder.

 The tissue distribution primarily in ovary tumors and breast cancer and, to a lesser extent, in normal lung and colon tumors indicates that the protein product of this gene is useful for the diagnosis and/or treatment of a variety of cancers, most notably cancers of the ovary, breast, or colon. Representative uses are described in the
25 "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, the expression of this gene product in a variety of cancers suggests that it may be a player in the progression of the disease, and may be a beneficial target for inhibitors as therapeutics. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or
30 receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may

show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1846 of SEQ ID NO:53, b is an integer of 15 to 1860, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

In an alternative reading frame, this gene shares sequence homology with a murine testosterone induced transcript (See Geneseq Accession No. 758299). This same region also shares sequence homology with a human cancer suppressor transfer factor protein (See Geneseq Accession No. R86875). The gene encoding the disclosed cDNA is thought to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

Preferred polypeptides of the invention comprise the following amino acid sequence:

KAPYSWLADSWPHPSRSPSAQEPRGSCCPSNPDPDDRYYNEAGISLYLAQTA
RGTAAPGEGPVYSTIDPAGEELQTFHGGFPQHPSGDLGPWSQYAPPEWSQG

(SEQ ID NO: 333). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in various embryonic/fetal tissues, particularly fetal brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, congenital birth defects, particularly of the central nervous system, and cancers, such as MEN. Similarly, polypeptides and antibodies directed to these

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal and embryonic tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of cancers, most notably cancers of the central nervous system, such as MEN, as well as the disorders of the central nervous system listed above. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, the expression within embryonic tissue and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the detection, treatment, and/or prevention of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Expression of this gene product in a variety of systems suggests that this gene may be a player in the progression of these diseases, and may be a beneficial target for inhibitors as therapeutics. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
5 formula of a-b, where a is any integer between 1 to 1649 of SEQ ID NO:54, b is an integer of 15 to 1663, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

10 The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1. This gene is highly homologous to bovine cytochrome b-5 reductase (See e.g., GENBANK: locus BOVCYB5R, accession M83104; Strittmatter et al., J. Biol. Chem. 267:2519-2523 (1992); the references
15 available through the accession number and the captioned reference are hereby incorporated herein by reference). Based on this homology, it is likely that this gene would have activity similar to NADH-cytochrome b5 reductase.

This gene is expressed primarily in liver and lung tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the liver and lung including chronic liver failure, bronchitis, emphysema, and chronic lung failure. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
25 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic and pulmonary systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hepatic, pulmonary, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid
30 and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 155 as residues: Arg-31 to Gln-37, Val-88 to Gly-95, Pro-110 to Gln-120, Gln-151 to Ala-163, Asp-231 to Trp-237, Pro-277 to Lys-287.

The tissue distribution in liver tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Representative uses are described in the "Hyperproliferative Disorders", "Infectious Disease", and "Binding Activity" sections below, in Example 11, and 27, and elsewhere herein. Alternatively, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of disorders associated with developing lungs, particularly in premature infants where the lungs are the last tissues to develop. The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of lung tumors, since the gene may be involved in the regulation of cell division, particularly since it is expressed in fetal tissue. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1618 of SEQ ID NO:55, b is an integer of 15 to 1632, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

This gene is expressed primarily in tonsil tissue and neutrophils, and, to a lesser extent, in testes tissue, brain and cerebellum tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the tonsils, immune system disorders, reproductive disorders, and neural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the tonsils, and the immune, reproductive, and neural systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, neural, reproductive, tonsils, cancerous and wounded tissues) or bodily fluids (e.g., lymph,
15 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
20 156 as residues: Pro-17 to Glu-26, Asp-60 to Val-72.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14,
25 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in tonsils as well as neutrophils suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other
30 processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

The tissue distribution in brain and cerebellum tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to

identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
10 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2219 of SEQ ID NO:56, b is an integer of 15 to 2233, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 47**

The translation product of this gene shares sequence homology with seven trans-membrane receptors and plectin, which is thought to be important in muscular dystrophy and multiple other diseases. The gene encoding the disclosed cDNA is thought to reside on chromosome 16. Accordingly, polynucleotides related to this
20 invention are useful as a marker in linkage analysis for chromosome 16.

This gene is expressed primarily in brain, fetal organs and placental tissue, and, to a lesser extent, in several other organs and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the central nervous system, fetal and developing organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
30 tissues or cells, particularly of the central nervous system, developing and fetal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, developing, cancerous

and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 157 as residues: Arg-13 to Trp-19, Leu-76 to Ala-92, Ser-100 to Arg-105.

The tissue distribution and homology to plectin and seven transmembrane receptors suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the central nervous
10 system, as well as developing and fetal systems. Moreover, the expression within fetal tissue indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, cancer, and other proliferative conditions.

Representative uses are described in the "Hyperproliferative Disorders" and
15 "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain
20 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
25 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in
30 proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1949 of SEQ ID NO:57, b is an integer of 15 to 1963, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

- Preferred polypeptides of the invention comprise the following amino acid sequence: LQQTMQAMLHFGGRLAQSLRGTSKEAASDPSDSPNLPTPGSWW (SEQ ID NO: 334),
20 EQLTQASRVYASGGTEGFPLSRWAPGRHGTAEEGAQERPLPTDE (SEQ ID NO: 335), MAPGRGLWLGRLFVPGGPAENENGALKSRRPSSWLPPTVSVLAL (SEQ ID NO: 336),
VKRGAPPEMPSPQELEASAPRMVQTHRAVRALCDHTAARPDQLS (SEQ ID NO: 337), FRRGEVLRVITTVDDEDWLRCGRDGMGLVPVGYTSLVL (SEQ ID
25 NO: 338), and/or
LQQTMQAMLHFGGRLAQSLRGTSKEAASDPSDSPNLPTPGSWWEQLTQASRVYASGGTEGFPLSRWAPGRHGTAEEGAQERPLPTDEMAPGRGLWLGRLFVPGGPAENENGALKSRRPSSWLPPTVSVLALVKRGAPPEMPSPQELEASAPRMVQTHRAVRALCDHTAARPDQLSFRRGEVLRVITTVDDEDWLRCGRDGMGLVPVGYTSLVL (SEQ ID NO: 339). Polynucleotides encoding these polypeptides
30 are also provided.

This gene is expressed primarily in synovium, synovial sarcoma, and chondrosarcoma tissues, and, to a lesser extent, in endometrial stromal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in skeletal tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g., arthritis, trauma, tendonitis, chondromalacia and inflammation). The protein product is useful in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e., spondyloepiphyseal dysplasia congenita, familial arthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Alternatively, the tissue distribution in endometrium suggests that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally.

Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelihood of implantation and development of a healthy

embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote healthy development of the endometrium. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1253 of SEQ ID NO:58, b is an integer of 15 to 1267, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

Preferred polypeptides of the invention comprise the following amino acid sequence:

ARACPRXGA AVEKLGGKPVQPSKPTCCSQVKA EGLIFAGLTGLKLLPSSLQ
20 RAVFVRQCLGFWNDGSRA LQ (SEQ ID NO: 340). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in hypothalamus and hepatocellular tumor and, to a lesser extent, in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, obesity, metabolic disorders, and hepatocellular tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
30 type(s). For a number of disorders of the above tissues or cells, particularly of the, endocrine system, hypothalamus and hepatocellular tumor, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell

types (e.g., hypothalamus, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in hypothalamus and hepatocellular tumors indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of obesity, metabolic disorders, and hepatocellular tumors. Similarly, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g., hyper-, hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1281 of SEQ ID NO:59, b is an integer of 15 to 1295, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

Preferred polypeptides of the invention comprise the following amino acid sequence: FQSVYHMKLQSSNLPASVYGNNLNCINSSSS (SEQ ID NO: 341). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain, placenta and breast.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, neurological and behavioural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS, immune and female reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, amniotic fluid, serum, plasma, urine, synovial fluid or cerebrospinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene in brain indicates that the protein products of this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. In addition, expression in breast and placenta suggests a role in the detection and/or treatment of female infertility and/or pregnancy disorders. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 901 of SEQ ID NO:60, b is an integer of 15 to 915, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 51

This gene is expressed primarily in adipocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, obesity, Nasu-Hakola disease, cardiovascular disease, non-insulin-dependent diabetes mellitus. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adipose, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 161 as residues: Asp-6 to Arg-12, Lys-31 to Leu-41.

The tissue distribution in adipose tissue suggests that the protein product of this gene is useful for the treatment and diagnosis of endocrine and metabolic disorders related to lipids and adipose tissue, such as obesity, Nasu-Hakola disease (membranous lipodystrophy), cardiovascular disease, lipidemia, non-insulin-dependent diabetes mellitus, stroke and carcinoma. Furthermore, the protein product of this gene may show utility in ameliorating conditions which occur secondary to aberrant fatty-acid metabolism (e.g., aberrant myelin sheath development), either directly or indirectly. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1431 of SEQ ID NO:61, b is an integer of 15 to 1445, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in
15 linkage analysis for chromosome 1.

This gene is expressed primarily in testes, endometrial tumor tissue, bone marrow and placenta tissue, and, to a lesser extent, in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive diseases and disorders, cancers and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of
25 the hematopoietic and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
30 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 162 as residues: Phe-32 to Gln-41, Gln-54 to Asn-68.

The tissue distribution in testes tissue and bone marrow suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the hematopoietic and reproductive systems, and cancers thereof. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g., endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1086 of SEQ ID NO:62, b is an

integer of 15 to 1100, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

5 The translation product of this gene has homology with metallothionine proteins from several organisms.

 This gene is expressed primarily in ovarian cancer, tonsils, and B-cell lymphoma.

 Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive defects, and lymphoid and ovarian cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
15 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and female reproductive systems, and of lymphoid and ovarian cancers, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid
20 and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 163 as residues: Leu-39 to Ser-47.

25 The tissue distribution in ovarian cancer, tonsils, and B-cell lymphoma suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, detection and/or treatment of female reproductive disorders, gonadal and general lymphoid neoplasias, and cancers thereof. Expression of this gene product in tonsils suggests a role in the regulation of the proliferation; survival;
30 differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of

cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1485 of SEQ ID NO:63, b is an integer of 15 to 1499, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

This gene is expressed primarily in adult kidney and pulmonary tissues, as well as in osteoblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, metabolic, endocrine and skeletal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine, skeletal, metabolic and
5 developmental systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine, skeletal, cancerous and wounded tissues) or bodily fluids (e.g., sputum, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
10 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 164 as residues: Ala-35 to Gly-45, Pro-67 to Pro-73, Pro-91 to Ser-97, Thr-127 to
Leu-139, Leu-143 to Asn-152, Ser-162 to Pro-167.

15 The tissue distribution in kidney tissue and osteoblasts suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and/or treatment of various endocrine and skeletal disorders. Furthermore, elevated levels of expression of this gene product in osteoblasts suggests that it may play a role in the survival, proliferation, and/or growth of osteoblasts. Therefore, it
20 may be useful in influencing bone mass in such conditions as osteoporosis. Alternatively, the tissue distribution in kidney suggests that this gene or gene product is useful in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria,
25 renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
30 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of $a-b$, where a is any integer between 1 to 641 of SEQ ID NO:64, b is an integer of 15 to 655, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:64, and where b is greater than or equal to $a + 14$.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

This gene is expressed primarily in neutrophils and embryonic tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders and cancers, and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
20 type(s). For a number of disorders of the above tissues or cells, particularly of the immune and developing systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell
25 sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 165 as residues: Gln-21 to Ala-33, Lys-48 to Pro-53.

30 The tissue distribution in neutrophils and embryonic tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, study and/or treatment of various developmental and immune system

disorders and cancers thereof, as well as cancers of other tissues where expression of this gene has been observed. Furthermore, expression within embryonic tissue and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the detection, treatment, and/or prevention of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy.

Alternatively, expression of this gene product in neutrophils also strongly suggests a role for this protein in immune function and immune surveillance. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1432 of SEQ ID NO:65, b is an integer of 15 to 1446, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

Preferred polypeptides of the invention comprise the following amino acid sequence: FDFIASLLKANRLSLQTCELLAAALLPSERYKAISI (SEQ ID NO: 342). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fetal liver, spleen and, to a lesser extent, in breast.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and haemopoietic diseases and/or disorders, in addition to, fetal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 166 as residues: Ile-50 to Ser-61, Pro-75 to Ser-104.

The tissue distribution in fetal liver and spleen suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of haemopoietic disorders involving stem cell production and maturation. Similarly, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may

also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

5 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 656 of SEQ ID NO:66, b is an integer of 15 to 670, where both a and b correspond to the positions of nucleotide

15 residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

This gene is expressed primarily in adult pulmonary cells.

Therefore, polynucleotides and polypeptides of the invention are useful as

20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, emphysema and other pulmonary diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell

25 type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., lung, cardiovascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, sputum, pulmonary surfactant, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or

30 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in adult pulmonary cells suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of disorders of the pulmonary systems, especially emphysema, asthma, and other similar dysfunctions. Representative uses are described elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1678 of SEQ ID NO:67, b is an integer of 15 to 1692, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

This gene is expressed primarily in hypothalamus (schizophrenic), and, to a lesser extent, in cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, schizophrenia and hypothalamic diseases and/or diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., CNS, cancerous and

wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution in hypothalamus (schizophrenic) and, to a lesser extent, in cerebellum suggests that the protein product of this gene is useful for detection, treatment, and/or prevention of neurological disorders, especially schizophrenia, neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative
10 Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction,
15 aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

20 Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
25 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 641 of SEQ ID NO:68, b is an integer of 15 to 655, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

This gene is expressed primarily in CD34 positive hematopoietic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic diseases and/or disorders; impaired immune function; susceptibility to infections; lymphomas and leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in CD34 positive cells indicates that the protein product of this gene is useful for the diagnosis and/or treatment of a variety of hematopoietic disorders. Expression of this gene product particularly in CD34 positive cells suggests that it plays a role in the proliferation; survival; differentiation; and/or activation of early stem and committed progenitor cells within the hematopoietic system. Thus, this gene product may be useful in determining the numbers and proportions of different hematopoietic cell lineages both in vitro and in vivo. Additionally, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages.

Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1604 of SEQ ID NO:69, b is an integer of 15 to 1618, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

Preferred polypeptides of the invention comprise the following amino acid sequence: IDLSFPSTNVSLEDRNTTKPSVNVG (SEQ ID NO: 343).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in dermatofibrosarcoma protuberance and 12 week old early human embryos.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, dermatofibrosarcoma; cancer; abnormal cell proliferation; embryological/developmental defects; inhibition of apoptosis; and hematopoietic diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin and epithelium, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., integumentary, reproductive, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this gene is useful for the diagnosis and/or treatment of abnormal cellular proliferation, such as cancer. Expression of this gene in dermatofibrosarcoma and 12 week early stage embryos indicates that it is involved in cellular proliferation and/or a block in differentiation. It may drive cellular proliferation directly, or it may play a role in inhibiting apoptosis or interfering with differentiation events. Similarly, this gene is useful for the treatment, diagnosis, and/or prevention of various skin disorders. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", "Infectious Disease", and "Regeneration" sections below, in Example 11, 19, and 20, and elsewhere herein. Briefly, the protein is useful in detecting, treating, and/or preventing congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e., keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e., wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e., lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus),

keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e., cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm).

5 Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders (i.e., arthritis, trauma, tendonitis, chondromalacia and inflammation, etc.), autoimmune disorders (i.e., rheumatoid arthritis, lupus, scleroderma, dermatomyositis, etc.), dwarfism, spinal deformation, joint abnormalities, and chondrodysplasias (i.e., spondyloepiphyseal
10 dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the
15 protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of
20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1788 of SEQ ID NO:70, b is an
25 integer of 15 to 1802, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

This gene is expressed primarily in neutrophils.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, disorders affecting the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system expression of this

5 gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual

10 not having the disorder.

The tissue distribution in neutrophils suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of immune system disorders, especially those affecting neutrophils. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below,

15 in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as

20 well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have

25 commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

30 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1278 of SEQ ID NO:71, b is an integer of 15 to 1292, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

Preferred polypeptides of the invention comprise the following amino acid sequence: LNILSLTVSSHCKL (SEQ ID NO: 344), INYHSGFIHQFLA (SEQ ID
15 NO: 345), and/or MANNSSLSSQFI (SEQ ID NO: 346). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in thymus tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
25 immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in
30 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 172 as residues: Pro-44 to Arg-50.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in thymus suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1040 of SEQ ID NO:72, b is an

integer of 15 to 1054, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

5 The translation product of this gene shares sequence homology with angiotensin II receptor which is thought to be important in ligand binding for blood pressure regulation. (See, e.g., GenBank Accession No. gil387891, gil1763532, and/or gil349736; all references available through these accessions are hereby incorporated herein by reference).

10 Preferred polypeptide fragments comprise the amino acid sequence (portion of extracellular domain):
PFWAAESALDFHWPFGGALCKMVLTVLNVYASIFLITALSVARY (SEQ ID NO: 347). Also preferred are the polynucleotides that encode this polypeptide fragment.

15 This gene is expressed primarily in 7TM-pbfd and PCMIX libraries (tissue types unknown).

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, blood pressure regulatory diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be
25 routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 173 as residues: Gln-88 to Ser-97.

The tissue distribution and homology to angiotensin II receptor suggests that the protein product of this gene is useful for the study, detection, treatment, and/or prevention of vascular diseases such as blood pressure regulatory disorders. Representative uses are described elsewhere herein. In particular, the extracellular region of the receptor can be used as a soluble antagonist. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are not limited to microvascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 719 of SEQ ID NO:73, b is an integer of 15 to 733, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

Preferred polypeptides of the invention comprise the following amino acid sequence: THADKNQVRNSN (SEQ ID NO: 348), QFLSWEQCTGNTESQ (SEQ ID NO: 349), VRRPKAKGXQTSN (SEQ ID NO: 350), PTQLNKHKPTTKERRRKGL (SEQ ID NO: 351), and/or LISKHENIY (SEQ ID NO: 352). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of immune system disorders, especially those affecting neutrophils. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
5 related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
10 formula of a-b, where a is any integer between 1 to 771 of SEQ ID NO:74, b is an integer of 15 to 785, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

15 Preferred polypeptides of the invention comprise the following amino acid sequence: TLYIXXMXTQTWRDQGRCGRDXINCIV (SEQ ID NO: 353). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain tissue from a manic depressive, in some cancer tissues such as ovarian cancer, and in spleen from a patient with chronic
20 lymphocytic leukemia and, to a lesser extent, in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain disorders (e.g., manic depression), and tumorigenesis. Similarly,
25 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system (CNS), reproductive system, and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain
30 tissues or cell types (e.g., brain, reproductive, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 175 as residues: Thr-29 to Ala-37, Arg-41 to Lys-46.

5 The tissue distribution primarily in brain tissue from a manic depressive indicates that the protein products of this gene are useful for diagnosing and treating manic depression and tumorigenesis.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
10 related to SEQ ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
15 formula of a-b, where a is any integer between 1 to 2327 of SEQ ID NO:75, b is an integer of 15 to 2341, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

20 Preferred polypeptides of the invention comprise the following amino acid sequence: SLCTPGRGWEESWGSSLPNLTGWSVSSLDNNDV (SEQ ID NO: 354). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in metastatic melanoma spleen, rhabdomyosarcoma, and IL-1 induced neutrophils and, to a lesser extent, in other
25 tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumorigenesis, metastasis and inflammatory disorders. Similarly,
30 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

skin, connective tissue and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skin, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in metastatic melanoma spleen, rhabdomyosarcoma, and IL-1 induced neutrophils indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of certain tumors such as melanoma, rhabdomyosarcoma and inflammatory disorders. Similarly, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (e.g., nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (e.g., keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (e.g., wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (e.g., lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (e.g., cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and immunotherapy targets for the above listed tumors and tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1868 of SEQ ID NO:76, b is an integer of 15 to 1882, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

Preferred polypeptides of the invention comprise the following amino acid sequence:

MQVALKEDLDALKEKFRTMESNQKSSFQEIPLNEELLSKQKQLEKIESGEM
10 GLNKVWINITEMNKQISLLTSVNHILKANVKSAA DLISLPTTVEGLQKSVASI
GXTLNSVHLAVEALQKTVDEHKKT MELLQSDMNQHFLKETPGSNQIIPSPSA
TSELDNKTHSENKQMGDRSATLKRQSLDQVTNRDTVQIKSIKKEG (SEQ
ID NO: 355), MQVALKEDLDALKEKFRTMESNQKSSFQEIPLNEELLSKQKQ
(SEQ ID NO: 356),
15 LEKIESGEMGLNKVWINITEMNKQISLLTSVNHILKANVKSAA (SEQ ID NO:
357), DLISLPTTVEGLQKSVASIGXTLNSVHLAVEALQKTVDEHKKT (SEQ ID
NO: 358), MELLQSDMNQHFLKETPGSNQIIPSPSATSELDNKTHSENKQ (SEQ
ID NO: 359), and/or MGDRSATLKRQSLDQVTNRDTVQIKSIKKEG (SEQ ID
NO: 360). Polynucleotides encoding these polypeptides are also provided.

20 This gene is expressed primarily in placental and infant brain tissues, and, to a lesser extent, in many normal and neoplastic cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, developmental disorders, cancer and general growth disorders.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, developing, and nervous systems, expression of this gene at
30 significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, developmental, neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
5 177 as residues: Cys-30 to Asn-44.

The tissue distribution in infant brain and embryonic tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, detection and/or treatment of growth and neoplastic disorders. Furthermore, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene
10 are useful for the detection, treatment, and/or prevention of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division. Embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein
15 may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Alternatively, the tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.

Representative uses are described in the "Regeneration" and
20 "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating
25 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated
30 expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal

differentiation or survival. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to
5 isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
10 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2878 of SEQ ID NO:77, b is an integer of 15 to 2892, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 68**

This gene is apparently exclusively in fetal heart tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, cardiovascular and growth defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing cardiovascular system, expression of this gene at significantly higher or lower levels may be routinely
30 detected in certain tissues or cell types (e.g., cardiovascular, heart, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such

a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal heart tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, detection and/or treatment of disorders and growth defects of heart development and function. Furthermore, the tissue distribution in fetal heart tissue indicates that the protein product of this gene is useful for the detection, treatment, and/or prevention of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stroke, angina, thrombosis, and wound healing. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1659 of SEQ ID NO:78, b is an integer of 15 to 1673, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

This gene is expressed primarily in pancreas islet cell tumor tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, digestive and metabolic defects and tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine, pancreas, cancerous and wounded tissues) or
5 bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in pancreas islet cell tumor tissue suggests that
10 polynucleotides and polypeptides corresponding to this gene are useful for the study, detection and/or treatment of hormonal and neoplastic disorders of endocrine organs and metabolism. Additionally, the tissue distribution indicates the protein product of this gene is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers. Representative uses are described in the "Biological Activity",
15 "Hyperproliferative Disorders", and "Binding Activity" sections below, in Example 11, 17, 18, 19, 20 and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of the Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g., hyper-,
20 hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothalamus, and testes. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
25 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1447 of SEQ ID NO:79, b is an integer of 15 to 1461, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, diseases and/or disorders of the tonsils, and disorders of the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the tonsils, and the immune system, expression of this gene at significantly higher or
15 lower levels may be routinely detected in certain tissues or cell types (e.g., tonsils, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
20 having the disorder.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of a variety of immune system disorders. Expression of this gene product in tonsils suggests a role in the regulation of the proliferation; survival; differentiation; and/or
25 activation of potentially all hematopoietic cell lineages, including blood stem cells. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the
30 treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker

and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have
5 commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
10 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of
15 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1503 of SEQ ID NO:80, b is an
20 integer of 15 to 1517, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

Preferred polypeptides of the invention comprise the following amino acid
25 sequence: SPQFLSSKSLPT (SEQ ID NO: 361). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain and spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, congenital brain disorders, including various forms of mental retardation, spina bifida, epilepsy, and various mood disorders, including bipolar and

unipolar depression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at
5 significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain, CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily
10 fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 181 as residues: Pro-42 to Lys-49, Lys-56 to Lys-71.

The tissue distribution in infant brain and spinal cord suggests that polynucleotides and polypeptides corresponding to this gene are useful for the
15 diagnosis and/or treatment of disorders of the brain and nervous system, including congenital brain disorders, including various forms of mental retardation, spina bifida, epilepsy, and various mood disorders, including bipolar and unipolar depression. It may also be useful in the treatment of such neurodegenerative disorders as schizophrenia; ALS; or Alzheimer's. Protein, as well as, antibodies directed against
20 the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of
25 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 560 of SEQ ID NO:81, b is an
30 integer of 15 to 574, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

Preferred polypeptides of the invention comprise the following amino acid sequence:

GPPSPRGLPSLPLHLAPRRYLQSR YACSQSSVSAAARRWGSGWMAWDPWN

5 QASGRYARITLLSVQACHQ

PTVWPRAGHSLPER YSLHPHNGDSTHLSGLLTVKCGA (SEQ ID NO: 362),

GPPSPRGLPSLPLHLAPRRYLQSR YACSQSSVSAAA (SEQ ID NO: 363),

RRWGSGWMAWDPWNQASGRYARITLLSVQACHQ (SEQ ID NO: 364), and/or

PTVWPRAGHSLPER YSLHPHNGDSTHLSGLLTVKCGA (SEQ ID NO: 365).

10 Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, infection, inflammation and other immune reactions or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in neutrophils indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of immune disorders, such as infection, inflammation, allergy and immunodeficiency. Therefore, this gene product may have clinical relevance in the treatment of impaired immunity, in the correction of autoimmunity, in immune modulation, in the treatment of allergy, and in the regulation of inflammation. It may also play a role in influencing differentiation of specific hematopoietic lineages, and may even affect the hematopoietic stem cell.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:82 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1441 of SEQ ID NO:82, b is an integer of 15 to 1455, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 73

Preferred polypeptides of the invention comprise the following amino acid sequence: NQENSLQTN SYLDSTESK (SEQ ID NO: 366). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in neutrophils and activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution neutrophils and T-cells indicates that the protein products of this gene are useful for diagnosis and treatment of immune related disorders including, infection, inflammation, allergy, tissue/organ transplantation, immunodeficiency, etc. Therefore, this gene product may have clinical relevance in the treatment of impaired immunity, in the correction of autoimmunity. in immune modulation, in the treatment of allergy, and in the regulation of inflammation. It may also play a role in influencing differentiation of specific hematopoietic lineages, and may even affect the hematopoietic stem cell. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1626 of SEQ ID NO:83, b is an integer of 15 to 1640, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:83, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

This gene is expressed primarily in hemangiopericytoma, placental tissue, and breast and endometrial tumor tissues, and, to a lesser extent, in various other normal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects and tumors of female reproductive organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
5 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in endometrial tumor tissue and placental tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the study, detection and/or treatment of reproductive system disorders and
10 neoplasias, as well as cancers of other tissues where expression of this gene has been observed. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
15 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 511 of SEQ ID NO:84, b is an integer of 15 to 525, where both a and b correspond to the positions of nucleotide
25 residues shown in SEQ ID NO:84, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

In an alternative reading frame, this gene shares homology with a DNA mismatch repair proteins, including PMS 4, and PMS1 (See Accession No. R95251,
30 gnllPIDId1008095 and pirIJC2399IJC2399).

Preferred amino acid fragments comprise the amino acid sequence:
QKRACFPFAFCRDCQFXEXSPAMLPVQPAXL (SEQ ID NO: 367),

VSAHGIWLFRRS (SEQ ID NO: 368), KHAAPPASLSLSLLLHHGQKR
ACFPFAFCRDCQFXEXSPAMLPVQPAXL (SEQ ID NO: 369). Polynucleotides
encoding these polypeptides are also provided.

This gene is expressed primarily in hematopoietic cells and tissues, such as
5 monocytes, primary dendritic cells, and thymus; and, to a lesser extent, in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, hematopoietic diseases and/or disorders; immune dysfunction;
10 susceptibility to infection; impaired immune surveillance; neurological disorders, and
cancers which may result from increased genetic instability. Similarly, polypeptides
and antibodies directed to these polypeptides are useful in providing immunological
probes for differential identification of the tissue(s) or cell type(s). For a number of
disorders of the above tissues or cells, particularly of the immune system, CNS, and
15 solid tissues, expression of this gene at significantly higher or lower levels may be
routinely detected in certain tissues or cell types (e.g., hematopoietic, cancerous and
wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid
and spinal fluid) or another tissue or cell sample taken from an individual having such
a disorder, relative to the standard gene expression level, i.e., the expression level in
20 healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution primarily in hematopoietic cells and tissues and the
homology to DNA mismatch repair proteins indicates that the protein product of this
gene is useful for the diagnosis and/or treatment of a variety of disorders, especially
cancer. Representative uses are described in the "Immune Activity" and "Infectious
25 Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere
herein. Briefly, the expression of this gene product in a number of hematopoietic cells
and tissues suggests that it may play a role in the proliferation; differentiation;
survival; and/or activation of a variety of hematopoietic lineages, particularly the
monocyte/macrophage pathway.

30 Expression of this gene product in a variety of brain tissues also suggests that
it may play a role in normal neuronal function or in establishment of neural
connectivity. Therefore, it may be useful in the treatment of neurological disorders,

such as Alzheimer's or Parkinson's. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the
5 protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:85 and may have been publicly available prior to conception of
10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 823 of SEQ ID NO:85, b is an
15 integer of 15 to 837, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 76

This gene is expressed primarily in T-cell lymphoma, endometrial tumors, and
20 infant brain cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T-cell lymphoma, endometrial tumor, and neurodegenerative or
25 developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, central nervous system, and reproductive systems, expression of this gene at significantly higher or lower levels
30 may be routinely detected in certain tissues or cell types (e.g., neural, immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken

from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
5 186 as residues: Glu-28 to Tyr-33, Gly-50 to Tyr-57.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for detecting and/or treating T-cell lymphoma, endometrial tumors, neurodegenerative or developmental disorders. The tissue distribution in infant brain cells suggests that polynucleotides and polypeptides
10 corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders.

Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection,
15 treatment, and/or prevention of Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment
20 and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may
25 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1560 of SEQ ID NO:86, b is an integer of 15 to 1574, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

This gene is expressed primarily in cancer cells, particular from hepatocellular carcinoma.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include; but are not limited to, hepatocellular carcinoma and other similar cancer, particularly of the liver. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
15 type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hepatic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such
20 a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tissues of cancerous origins, such as hepatocellular carcinoma tissue, suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of cancers, most
25 notably cancers of the liver, such as hepatocellular carcinoma. Expression of this gene product in a variety of cancers suggests that this gene may be a player in the progression of these diseases, and may be a beneficial target for inhibitors as therapeutics. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:87 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
5 formula of a-b, where a is any integer between 1 to 1614 of SEQ ID NO:87, b is an integer of 15 to 1628, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:87, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

10 This gene is expressed primarily in T-cell lymphoma, and, to a lesser extent, in hepatocellular tumor tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to; T-cell lymphoma, hepatocellular tumors, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hepatic systems, expression of this gene at significantly higher or lower
20 levels may be routinely detected in certain tissues or cell types (e.g., immune, hepatic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the
25 disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 188 as residues: Pro-46 to Asn-58.

The tissue distribution in T-cell lymphoma and hepatocellular tumor tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful
30 for the detection and/or treatment of T-cell lymphomas and hepatocellular tumors, as well as cancers of other tissues where expression of this gene has been observed. Representative uses are described in the "Immune Activity" and "Infectious Disease"

sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

- 5 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:88 and may have been publicly available prior to conception of
10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1781 of SEQ ID NO:88, b is an
15 integer of 15 to 1795, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:88, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 79

This gene is expressed primarily in brain tissue, and, to a lesser extent, in
20 ntera2 cell lines, melanocytes, normal colon, and T-helper cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases and/or conditions. Similarly, polypeptides
25 and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, immune, hematopoietic, gastrointestinal, and
30 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 189 as residues: Met-1 to Trp-6.

5 The tissue distribution in brain tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for detecting and/or treating neurodegenerative diseases of the central nervous system. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the tissue distribution
10 suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered
15 behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or
20 receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1850 of SEQ ID NO:89, b is an

integer of 15 to 1864, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

5 The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

 Preferred polypeptides of the invention comprise the following amino acid sequence: IPEEASCFPSAV (SEQ ID NO: 370), EILFGKLKSKAALCTQG (SEQ ID
10 NO: 371), HADRYTCCRCLSPFSLAGL (SEQ ID NO: 372), LSDPLLLPDCSFSFN (SEQ ID NO: 373), KAVAYANVSCRRFKHKTTKLGPIQW (SEQ ID NO: 374), PSSQSPEPPQPLSLFVTRLPNLYDFP (SEQ ID NO: 375), and/or SRQIICTNLCKCTPICFLF (SEQ ID NO: 376). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed primarily in breast tissue, fetal liver and adult hepatoma tissues, and, to a lesser extent, in merkel cells and osteoblasts.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, cancers of the liver or breast. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the glandular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain
25 tissues or cell types (e.g., breast, liver, cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 190 as residues: Asn-25 to Gln-50.

The tissue distribution in breast and hepatoma tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating tumors of the breast or liver. Furthermore, the expression in the breast tissue may indicate its uses in breast neoplasia and breast cancers, such as

- 5 fibroadenoma, papillary carcinoma, ductal carcinoma, Paget's disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases.

- Alternatively, the tissue distribution suggests that polynucleotides and
10 polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and immunotherapy targets for the above listed tumors and
15 tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1969 of SEQ ID NO:90, b is an integer of 15 to 1983, where both a and b correspond to the positions of nucleotide
25 residues shown in SEQ ID NO:90, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

This gene is expressed primarily in thymus and brain tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the immune system and diseases of the

brain, including various types of mood disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in thymus suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution in brain tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette

Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of

5 developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show

10 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1943 of SEQ ID NO:91, b is an integer of 15 to 1957, where both a and b correspond to the positions of nucleotide

20 residues shown in SEQ ID NO:91, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

Preferred polypeptides of the invention comprise the following amino acid sequences: MLLPVNTLLYI (SEQ ID NO: 377), LLTPLCFFYGTSRP (SEQ ID NO:

25 378), PYLELVT (SEQ ID NO: 379), LLKKKKQSVGFVS (SEQ ID NO: 380), CILEAGR (SEQ ID NO: 381), MGFSAPTPGPL (SEQ ID NO: 382), FDLRRLLSIV (SEQ ID NO: 383), AFCPHVTPCKYAVIHTV (SEQ ID NO: 384), NTPLLFLWDLQ (SEQ ID NO: 385), ATIFRTSYLIKKEKTV (SEQ ID NO: 386), WLLSLHLGGREVRAGAP (SEQ ID NO: 387), QTLQEGSLHSI (SEQ ID NO:

30 388), and/or

MGFSAPTPGPLFDLRRLLSIVAFCPHVTPCKYAVIHTVNTPLLFLWDLQATIF

RTSYLIKKEKTVCWLLSLHLGGREVRACAPQTLQEGSLHSI (SEQ ID NO: 389). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain and breast tissues, and, to a lesser extent, in several other cell and tissue types including colon and liver tissues.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast and brain cancers, mood disorders, dementia, and Alzheimer's disease. Similarly, polypeptides and antibodies directed to these polypeptides are
10 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and lactations systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph,
15 breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
20 192 as residues: Gly-21 to Tyr-27.

The expression of this gene in breast tissue may indicate its uses in breast neoplasia and breast cancers, such as fibroadenoma, papillary carcinoma, ductal carcinoma, Paget's disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile
25 hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Alternatively, the tissue distribution of this gene in brain tissue suggests that the translation product of this gene is useful for the detection and/or
30 treatment of brain cancers and neural disorders, such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning

disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 559 of SEQ ID NO:92, b is an integer of 15 to 573, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 83

This gene is expressed primarily in liver and, to a lesser extent, in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, liver/hepatocyte disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell

types (e.g., liver, cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in liver indicates that the protein products of this gene are useful for detection, treatment, and/or prevention of liver (hepatocyte) disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells).

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and immunotherapy targets for the above listed tumors and tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:93 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1198 of SEQ ID NO:93, b is an integer of 15 to 1212, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:93, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 84

Preferred polypeptides of the invention comprise the following amino acid sequence: YWVSISQRSVCQQARTSIFFKDGLSREKYSNNG (SEQ ID NO: 390). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, including AIDS and various other diseases in which the immune system is suppressed. Similarly, polypeptides and antibodies directed to

these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution in T cells indicates that the polypeptides or polynucleotides are useful for treatment, prophylaxis, and diagnosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. The polypeptides or polynucleotides of the present invention are also useful in the treatment, prophylaxis, and detection of thymus disorders, such as Grave's Disease, lymphocytic thyroiditis, hyperthyroidism, and hypothyroidism. Similarly, elevated levels of expression of this gene product in T cell lineages suggests that it may play an active role in normal T cell function and in the regulation of the immune response. For example, this gene product may be involved in T cell activation, in the activation or control of differentiation of other hematopoietic cell lineages, in antigen recognition, or in T cell proliferation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:94 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1130 of SEQ ID NO:94, b is an integer of 15 to 1144, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:94, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

The translation product of this gene shares sequence homology with a protein which was found to accumulate during growth-factor-induced proliferation and transformation of normal rat fibroblasts (See, Glaichenhaus, N., and Cuzin, F., Cell 50:1081 (1987); and Genbank Acc. No. gil207250).

Preferred polypeptides of the invention comprise the following amino acid sequence:

LSVRAPGVPAARPLSSARQAGAGRGELRGQRLWLGPECGCGAGQAGSMLR
10 AVGSLLRLGRGLTVRCGPGAPLEATRRPAPALPPRGLPCYSSGGAPSNSGPQG
HGEIHRVPTQRRPSQFDKKILLWTGRFKSMEEIPPRIPPEMIDTARNKARVKAC
YI (SEQ ID NO: 391), LSVRAPGVPAARPLSSARQAGAGRGELRGQRLWLG
(SEQ ID NO: 392), PECGCGAGQAGSMLRAVGSLLRLGRGLTVRCGPG (SEQ
ID NO: 393), APLEATRRPAPALPPRGLPCYSSGGAPSNSGPQG (SEQ ID NO:
15 394), HGEIHRVPTQRRPSQFDKKILLWTGRF (SEQ ID NO: 395), and/or
KSMEEIPPRIPPEMIDTARNKARVKACYI (SEQ ID NO: 396). Polynucleotides
encoding these polypeptides are also provided.

This gene is expressed primarily in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, developmental anomalies or fetal deficiencies, cancers or neoplastic
conditions. Similarly, polypeptides and antibodies directed to these polypeptides are
useful in providing immunological probes for differential identification of the tissue(s)
25 or cell type(s). For a number of disorders of the above tissues or cells, particularly of
the developing embryo, expression of this gene at significantly higher or lower levels
may be routinely detected in certain tissues or cell types (e.g., embryonic, placental,
cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum,
plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
30 from an individual having such a disorder, relative to the standard gene expression
level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
having the disorder.

The tissue distribution and homology to a protein which was found to accumulate during proliferation and transformation of normal fibroblasts suggests that the protein product of this gene is useful for the treatment and diagnosis of developmental anomalies or fetal deficiencies, neoplasms and cancers. Additionally, the tissue distribution in placenta suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta suggests that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also suggests that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:95 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1260 of SEQ ID NO:95, b is an integer of 15 to 1274, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:95, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 86

The gene encoding the disclosed cDNA is thought to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

5 This gene is expressed primarily in T-cell lymphoma and synovial sarcoma tissues, and, to a lesser extent, in fetal liver/spleen tissue and synovial fibroblasts.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, T-Cell lymphoma and synovial sarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain
15 tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 196 as residues: Gly-4 to His-10, Asp-32 to Val-38.

 The tissue distribution in T-cell lymphoma and synovial sarcoma tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of T-cell lymphomas and synovial sarcomas, as
25 well as cancers of other tissues where expression of this gene has been observed. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents
30 that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:96 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1766 of SEQ ID NO:96, b is an integer of 15 to 1780, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:96, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 87

The gene encoding the disclosed cDNA is believed to reside on chromosome 10. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 10.

This gene is expressed primarily in brain and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, kidney diseases and various diseases of the brain including mood disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., kidney, CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or cerebrospinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 197 as residues: Arg-68 to Lys-78.

The tissue distribution in kidney suggests that this gene or gene product is useful in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Alternatively, the tissue distribution in brain suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the brain and nervous system. It may be useful in the treatment of such neurodegenerative disorders as schizophrenia, ALS, or Alzheimer's. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:97 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2051 of SEQ ID NO:97, b is an integer of 15 to 2065, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:97, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 88

It has been discovered that this gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and inflammatory systems,

expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
5 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 198 as residues: Pro-41 to Gln-48.

The tissue distribution in neutrophils indicates that the protein products of this
10 gene are useful for the study, diagnosis and/or treatment of immune and inflammatory diseases. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell
15 lineages, including blood stem cells. Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as
20 well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have
25 commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
30 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:98 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1140 of SEQ ID NO:98, b is an integer of 15 to 1154, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:98, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 89

Preferred polypeptides of the invention comprise the following amino acid sequence: ELAIGESCS (SEQ ID NO: 397). Polynucleotides encoding these
15 polypeptides are also provided.

This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, developmental, degenerative and behavioral diseases of the brain such as schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, transmissible spongiform encephalopathies (TSE), Creutzfeldt-Jakob disease (CJD), specific brain tumors, aphasia, mania, depression, dementia, paranoia, addictive behavior and sleep disorders. Similarly, polypeptides and antibodies directed to these
25 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
30 synovial fluid or cerebrospinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 199 as residues: Gly-45 to Thr-50.

5 The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly,
10 the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive
15 compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

 In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is
20 involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the
25 protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:99 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 601 of SEQ ID NO:99, b is an integer of 15 to 615, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:99, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 90

This gene is expressed primarily in brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and behavioural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., CNS, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or cerebrospinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain indicates that the protein products of this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:100 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1610 of SEQ ID NO:100, b is an integer of 15 to 1624, where both a and b correspond to the positions of nucleotide
5 residues shown in SEQ ID NO:100, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 91

Preferred polypeptides of the invention comprise the following amino acid sequence: PVIWPDGKRIVLLAEVS (SEQ ID NO: 398). Polynucleotides encoding
10 these polypeptides are also provided.

This gene is expressed primarily in adrenal gland tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, adrenal gland cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adrenal system, expression of this gene at
20 types (e.g., adrenal gland, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 201 as residues: Arg-49 to Gln-56.

The tissue distribution in adrenal gland indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders involving the adrenal gland. Expression of this gene product in adrenal
30 gland tumor indicates that it may play a role in the proliferation of cells of the adrenal gland, or potentially in the proliferation of cells in general. In such an event, it may play a role in determining the course and severity of cancer. Alternatively, it may play

a role in the normal function of adrenal glands, such as in the production of corticosteroids, androgens, or epinephrines. Thus it may play a role in general homeostasis, as well as in disorders involving the androgen hormones. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker
5 and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:101 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1742 of SEQ ID NO:101, b is an integer of 15 to 1756, where both a and b correspond to the positions of nucleotide
15 residues shown in SEQ ID NO:101, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 92

The gene encoding the disclosed cDNA is thought to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in
20 linkage analysis for chromosome 2.

This gene is expressed in multiple tissues, including the thymus, and cell types, including B cells and monocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders and/or disorders afflicting the immune system, such as AIDS and autoimmune diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
30 tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph,

serum, plasma, urine, synovial fluid and spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune system tissues and cells suggests that

5 polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders affecting the immune system, especially autoimmune diseases and AIDS. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, this gene product may be involved in the

10 regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker

15 and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of

20 various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show

25 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:102 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1402 of SEQ ID NO:102, b is an integer of 15 to 1416, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:102, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 93

This gene is expressed primarily in fetal lung tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, lung diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lung, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell
15 types (e.g., pulmonary, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, sputum, pulmonary surfactant, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 203 as residues: Leu-32 to His-38.

The tissue distribution in fetal lung tissue suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of lung diseases and/or disorders. Representative uses are described elsewhere herein.
25 Furthermore, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of disorders associated with developing lungs, particularly in premature infants where the lungs are the last tissues to develop. The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and
30 intervention of lung tumors, since the gene may be involved in the regulation of cell division, particularly since it is expressed in fetal tissue. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to

isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:103 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 690 of SEQ ID NO:103, b is an integer of 15 to 704, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:103, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 94

- Preferred polypeptides of the invention comprise the following amino acid sequence: CFLSVSFQWN (SEQ ID NO: 399), VTIAQVGIFVCFVHCCT (SEQ ID NO: 400), PGQVPSKHLGSNASVRA (SEQ ID NO: 401),
20 DEGAKVQRRPWGSQTHSPVLFL (SEQ ID NO: 402), LTRPGLWGSLLPVQQQRG (SEQ ID NO: 403), CASLGVLNRANRSPCV (SEQ ID NO: 404), SWLEVTTLSAPGPVITTY (SEQ ID NO: 405), PGQWVREIXLVGRAVARV (SEQ ID NO: 406), LTWPPXGPMGTWVPGF (SEQ ID NO: 407), MADIPGTFLALGCHGQR (SEQ ID NO: 408),
25 VGRGSWASGWTNQSA (SEQ ID NO: 409), and/or PDHPLPVGLLEAWRVE (SEQ ID NO: 410). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily neutrophils and eosinophils, and, to a lesser extent, in bone marrow and fetal liver/spleen tissue.

- Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, asthma and diseases and/or disorders afflicting the immune system.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be
5 routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 204 as residues: Ser-2 to Trp-7.

The tissue distribution in immune system cells and tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of asthma or other disorders affecting the immune system.
15 Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

20 Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory
25 bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify
30 agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:104 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1245 of SEQ ID NO:104, b is an integer of 15 to 1259, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:104, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 95

This gene shares sequence homology to the rat cornichon-like protein (See Genbank Accession No. 2317276), the murine cornichon protein (See Genbank Accession No. gil2460430), and the human cornichon protein (See Genbank Accession No. gil4063709). The Drosophila cornichon gene is thought to be involved in signaling processes necessary for both anterior-posterior and dorsal-ventral pattern formation in Drosophila. Thus, it is likely that this gene plays a similar role in human development.

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in endometrial tumor tissue and infant brain tissue, and, to a lesser extent, in frontal cortex tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometrial tumor, and neural and developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neural and reproductive organs, expression of this gene at

significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
5 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 205 as residues: Glu-33 to Phe-38.

The tissue distribution in infant brain tissue and frontal cortex tissue, and the
10 homology to cornichon proteins, suggests that polynucleotides and polypeptides corresponding to this gene are useful for detecting and/or treating neural and developmental disorders. The tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers
15 Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of
20 developmental disorders associated with the developing embryo, or sexually-linked disorders.

Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the elevated expression of this gene product within the
25 frontal cortex of the brain suggests that it may be involved in neuronal survival; synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. Alternatively, the tissue distribution in endometrial tumor tissue suggests that the translation product of this gene is useful for the detection and/or treatment of endometrial tumors and/or
30 reproductive disorders, as well as tumors of other tissues where expression of this gene has been observed. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or

receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:105 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1790 of SEQ ID NO:105, b is an integer of 15 to 1804, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:105, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 96

The translation product of this gene shares significant sequence homology with a protein which was recently sequenced by another group, which was named paraplegin by this group (See Genbank Accession No. g3273089). The gene encoding
20 the disclosed cDNA is thought to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

Preferred polypeptides of the invention comprise the following amino acid sequence:

25 LARADPPGCRRRGWRPSSAELQLRLLTPTFEGINGLLKQHLVQNPVRLWQL
LGGTFYFNTSRLKQKNKE KDKSKGKAPEEDEXERRRRERDDQ (SEQ ID NO:
411). Polynucleotides encoding these polypeptides are also provided.

When tested against Jurkat T-cell cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates
30 T-cells, and to a lesser extent other immune cells, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The

Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

- 5 This gene is expressed primarily in Jurkat T-cells, Macrophage, T-Cell Lymphoma, tonsils, and salivary glands.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T-Cell lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 206 as residues: Met-1 to Leu-6, Asp-84 to Lys-89, Asp-124 to Gly-130, Ser-138 to Trp-143, His-145 to Ser-153, Thr-170 to Pro-183, Trp-191 to Pro-198.

 The tissue distribution in immune tissues and T-cells, in conjunction with the detected GAS biological activity data, suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of T-cell lymphomas. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in T cell lymphoma suggests that it may play a role in the proliferation of the lymphoid cell lineages, and may be involved in normal antigen recognition and activation of T cells during the immune process. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors,

to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:106 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
10 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 957 of SEQ ID NO:106, b is an integer of 15 to 971, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:106, and where b is greater than or equal to a + 14.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 97**

Preferred polypeptides of the invention comprise the following amino acid sequence: FLRFWCTCHVSS (SEQ ID NO: 412). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in bladder.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the bladder, including bladder cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
25 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., bladder, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or
30 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in bladder indicates that the polynucleotides and polypeptides corresponding to this gene are useful for treatment and/or diagnosis of urinary tract disorders (e.g., cystitis, urinary tract calculi, incontinence) and bladder tumors or cancers. Protein, as well as, antibodies directed against the protein may
5 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:107 and may have been publicly available prior to conception
10 of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 807 of SEQ ID NO:107, b is an
15 integer of 15 to 821, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:107, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HKABZ65	209683 03/20/98	pCMV/Sport 2.0	11	1191	1	1191	69	69	111	1	17	18	243
2	HNGIC80	209683 03/20/98	Uni-ZAP XR	12	1251	1	1251	24	24	112	1	24	25	41
3	HDPUG50	209745 04/07/98	pCMV/Sport 3.0	13	1734	1	1734	22	22	113	1	34	35	526
4	HAEAB66	209745 04/07/98	pBluescript SK-	14	1540	914	1537	105	105	114	1	30	31	354
5	HHEPF59	209746 04/07/98	pCMV/Sport 3.0	15	1558	1	1558	38	38	115	1	21	22	63
6	HE9BK23	209683 03/20/98	Uni-ZAP XR	16	1636	1	1636	39	39	116	1	21	22	309
7	HCYBI36	209683 03/20/98	pBluescript SK-	17	1256	148	1256	235	235	117	1	23	24	211
8	HSSDX51	209683 03/20/98	Uni-ZAP XR	18	1143	1	1143	133	133	118	1	20	21	50
9	HSDAJ46	209746 04/07/98	Uni-ZAP XR	19	1537	92	1537	299	299	119	1	18	19	262
10	HRACG45	209745 04/07/98	pCMV/Sport 3.0	20	2672	222	2672	178	178	120	1	42	43	270
11	HAPPW30	209683 03/20/98	Uni-ZAP XR	21	1508	14	1501	54	54	121	1	22	23	91

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
12	HE2ES51	209745 04/07/98	Uni-ZAP XR	22	1447	1	1447	77	77	122	1	14	15	222
13	HTXDW56	209746 04/07/98	Uni-ZAP XR	23	1583	1	1583	217	217	123	1	22	23	201
14	HEEAG23	209745 04/07/98	Uni-ZAP XR	24	1669	25	1280	57	57	124	1	18	19	46
15	HDPKI93	209745 04/07/98	pCMV/Sport 3.0	25	1053	1	1053	46	46	125	1	21	22	305
16	HDLAC10	209745 04/07/98	pCMV/Sport 2.0	26	1477	1	1477	132	132	126	1	29	30	81
17	HDPOH06	209745 04/07/98	pCMV/Sport 3.0	27	2504	1	2504	252	252	127	1	29	30	242
18	HCE4G61	209745 04/07/98	Uni-ZAP XR	28	1866	1	1866	130	130	128	1	23	24	285
18	HCE4G61	209745 04/07/98	Uni-ZAP XR	108	1779	1	1720	125	125	208	1	20	21	81
19	HCWUI13	209745 04/07/98	ZAP Express	29	1501	1	1501	80	80	129	1	18	19	157
20	HDPSP01	209745 04/07/98	pCMV/Sport 3.0	30	1752	1	1752	227	227	130	1	20	21	308
21	HHPEN62	209746 04/07/98	Uni-ZAP XR	31	2152	141	2152	183	183	131	1	27	28	508

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	S' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
22	HUKBT29	209746 04/07/98	Lambda ZAP II	32	1757	56	1757	74	74	132	1	19	20	506
23	HMAJR50	209683 03/20/98	Uni-ZAP XR	33	1466	32	1466	70	70	133	1	21	22	48
24	HBIMB51	209683 03/20/98	pCMV Sport 3.0	34	526	1	526	93	93	134	1	21	22	130
25	HE8DX88	209683 03/20/98	Uni-ZAP XR	35	2412	1	2412	256	256	135	1	29	30	43
26	HNGHT03	209746 04/07/98	Uni-ZAP XR	36	1274	65	1274	305	305	136	1	24	25	91
27	HWABU17	209745 04/07/98	pCMV Sport 3.0	37	1036	1	1036	202	202	137	1	18	19	266
28	HDTAT90	209746 04/07/98	pCMV Sport 2.0	38	1379	8	1379	78	78	138	1	26	27	434
29	HHFGR93	209746 04/07/98	Uni-ZAP XR	39	1932	1	1836	130	130	139	1	29	30	236
30	HOVCB25	209746 04/07/98	pSport1	40	1430	1	1430	150	150	140	1	18	19	99
31	HSYAV66	209746 04/07/98	pCMV Sport 3.0	41	1407	1	1407	186	186	141	1	28	29	69
32	HFPCT29	209683 03/20/98	Uni-ZAP XR	42	950	1	950	268	268	142	1	26	27	61

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
33	HAWAT25	209683 03/20/98	pBluescript SK-	43	1004	56	1004	149	149	143	1	32	33	88
34	HNHFR04	209683 03/20/98	Uni-ZAP XR	44	1681	1	1681	71	71	144	1	21	22	78
35	HOSFT61	209683 03/20/98	Uni-ZAP XR	45	1365	1	1365	211	211	145	1	21	22	90
36	HBJTO81	209683 03/20/98	Uni-ZAP XR	46	1137	1	1137	220	220	146	1	23	24	68
37	HADCL55	209745 04/07/98	pSport1	47	2763	15	2763	60	60	147	1	29	30	43
38	HAGGJ80	209745 04/07/98	Uni-ZAP XR	48	1576	1	1576	40	40	148	1	34	35	84
39	HAIBO81	209745 04/07/98	Uni-ZAP XR	49	1348	1	1348	250	250	149	1	18	19	63
40	HBBBC37	209745 04/07/98	pCMVSPORT 1	50	1264	1	1264	81	81	150	1	17	18	61
41	HBJMX85	209745 04/07/98	Uni-ZAP XR	51	1660	39	1660	45	45	151	1	18	19	82
42	HCEES66	209745 04/07/98	Uni-ZAP XR	52	1678	1	1678	178	178	152	1	39	40	46
43	HCEMP62	209745 04/07/98	Uni-ZAP XR	53	1860	269	1726	352	352	153	1	30	31	187

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
44	HE2FB90	209746 04/07/98	Uni-ZAP XR	54	1663	1	1663	205	205	154	1	27	28	113
45	HTHDJ94	209746 04/07/98	Uni-ZAP XR	55	1632	20	1632	66	66	155	1	26	27	292
46	HTOHJ89	209746 04/07/98	Uni-ZAP XR	56	2233	1	2233	42	42	156	1	17	18	86
47	HUSHB62	209745 04/07/98	Lambda ZAP II	57	1963	1	1760	130	130	157	1	49	50	106
48	HSXAG02	209683 03/20/98	Uni-ZAP XR	58	1267	411	1243	600	600	158	1	22	23	58
49	HHTLH52	209683 03/20/98	ZAP Express	59	1295	1	1295	218	218	159	1	22	23	40
50	HCFMS95	209683 03/20/98	pSport1	60	915	1	915	123	123	160	1	22	23	65
51	HOUCT90	209683 03/20/98	Uni-ZAP XR	61	1445	1	1445	74	74	161	1	30	31	46
52	HCFLR78	209745 04/07/98	pSport1	62	1100	224	1100	475	475	162	1	16	17	140
53	HTOHT18	209745 04/07/98	Uni-ZAP XR	63	1499	267	1499	433	433	163	1	24	25	53
54	HKPMB11	209745 04/07/98	pBluescript	64	655	1	655	55	55	164	1	25	26	167

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
55	HNFHS3	209745 04/07/98	Uni-ZAP XR	65	1446	1	1446	171	171	165	1	18	19	62
56	HAIBU10	209745 04/07/98	Uni-ZAP XR	66	670	1	669	201	201	166	1	20	21	113
57	HAPOK30	209745 04/07/98	Uni-ZAP XR	67	1692	1	1692	300	300	167	1	19	20	61
58	HCEEM18	209745 04/07/98	Uni-ZAP XR	68	655	18	655	157	157	168	1	30	31	41
59	HCWUA22	209745 04/07/98	ZAP Express	69	1618	48	1618	233	233	169	1	33	34	42
60	HDSAG91	209745 04/07/98	Uni-ZAP XR	70	1802	1	1802	156	156	170	1	23	24	47
61	HNEDJ35	209746 04/07/98	Uni-ZAP XR	71	1292	1	1292	71	71	171	1	36	37	50
62	HTHBH29	209746 04/07/98	Uni-ZAP XR	72	1054	1	1054	52	52	172	1	24	25	56
63	H7TBA62	209745 04/07/98	PCRII	73	733	9	718	224	224	173	1	36	37	170
64	HNGIO50	209746 04/07/98	Uni-ZAP XR	74	785	1	785	132	132	174	1	27	28	44
65	HMIW81	209683 03/20/98	Uni-ZAP XR	75	2341	1	2215	229	229	175	1	17	18	46

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
66	HMMC160	209683 03/20/98	pSport1	76	1882	1	1882	132	132	176	1	16	17	41
67	HDP1009	209745 04/07/98	pCMV/Sport 3.0	77	2892	17	2892	85	85	177	1	36	37	47
68	HHFHH34	209745 04/07/98	Uni-ZAP XR	78	1673	1	1673	16	16	178	1	22	23	70
69	HISCL83	209745 04/07/98	pSport1	79	1461	1	1461	259	259	179	1	21	22	41
70	HTOAI70	209746 04/07/98	Uni-ZAP XR	80	1517	1	1517	190	190	180	1	19	20	92
70	HTOAI70	209746 04/07/98	Uni-ZAP XR	109	1518	1	1518	190	190	209	1	19	20	42
71	HSDER95	209683 03/20/98	Uni-ZAP XR	81	574	1	574	72	72	181	1	25	26	71
72	HNECL25	209683 03/20/98	Uni-ZAP XR	82	1455	1	1455	322	322	182	1	32	33	66
73	HNFGZ45	209683 03/20/98	Uni-ZAP XR	83	1640	1	1640	450	450	183	1	38	39	70
74	HHGCU49	209745 04/07/98	Lambda ZAP II	84	525	1	525	173	173	184	1	23	24	40
75	HDPND68	209745 04/07/98	pCMV/Sport 3.0	85	837	1	837	154	154	185	1	17	18	66

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
76	HETDT81	209746 04/07/98	Uni-ZAP XR	86	1574	1	1574	189	189	186	1	25	26	66
77	HHLBA14	209746 04/07/98	pBluescript SK-	87	1628	353	1627	546	546	187	1	24	25	48
78	HLTBU43	209746 04/07/98	Uni-ZAP XR	88	1795	1	1795	198	198	188	1	19	20	66
79	HNTSJ84	209746 04/07/98	pSport1	89	1864	239	1864	336	336	189	1	22	23	57
80	HOHCG16	209746 04/07/98	pCMVSPORT 2.0	90	1983	1	1983	257	257	190	1	18	19	52
81	HTHCB31	209746 04/07/98	Uni-ZAP XR	91	1957	1	1957	46	46	191	1	17	18	43
82	HUKAM16	209746 04/07/98	Lambda ZAP II	92	573	1	573	178	178	192	1	23	24	52
83	HLDOIC5	209683 03/20/98	pCMVSPORT 3.0	93	1212	1	1212	313	313	193	1	20	21	40
84	HTXKFI0	209683 03/20/98	Uni-ZAP XR	94	1144	1	1144	334	334	194	1	32	33	71
85	HPMAI22	209683 03/20/98	Uni-ZAP XR	95	1274	334	1274	483	483	195	1	16	17	59
86	HL2AG57	209746 04/07/98	Uni-ZAP XR	96	1780	349	1780	560	560	196	1	31	32	80

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
87	HUSAM59	209683 03/20/98	Lambda ZAP II	97	2065	1	2065	475	475	197	1	17	18	78
88	HNGGR26	209745 04/07/98	Uni-ZAP XR	98	1154	1	1154	50	50	198	1	27	28	115
89	HTLCX30	209683 03/20/98	Uni-ZAP XR	99	615	1	459	60	60	199	1	28	29	50
90	HCEBC87	209683 03/20/98	Uni-ZAP XR	100	1624	243	1624	517	517	200	1	23	24	57
91	HATCB92	209683 03/20/98	Uni-ZAP XR	101	1756	1	1756	247	247	201	1	40	41	56
92	HMSCX69	209746 04/07/98	Uni-ZAP XR	102	1416	207	1416	246	246	202	1	16	17	49
93	HLHAL68	209746 04/07/98	Uni-ZAP XR	103	704	1	704	30	30	203	1	21	22	44
94	HEOMR73	209746 04/07/98	pSport1	104	1259	644	1259	354	354	204	1	24	25	44
95	HETIB83	209746 04/07/98	Uni-ZAP XR	105	1804	1	1804	104	104	205	1	30	31	160
96	HJPDD28	209746 04/07/98	Uni-ZAP XR	106	971	260	971	283	283	206	1	21	22	198
96	HJPDD28	209746 04/07/98	Uni-ZAP XR	110	921	1	921	31	31	210	1	21	22	96

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
97	HBAMB15	209683 03/20/98	pSport1	107	821	330	821	390	390	207	1	19	20	59

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization

probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID
5 NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or
10 deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

15 Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC,
20 as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a
25 suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed
30 herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

5 The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

10 The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during
15 recombinant production.

 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40
20 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

25 Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the
30 information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of

these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95%

"identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence

that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

10 For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

25 By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid.

30

These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

- 5 As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity.
- 10 Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.
- 15 If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from
- 20 the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and
- 25
- 30

C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

- 5 For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

- 20 The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

30 Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.

Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

5 Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988
10 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., *J. Biotechnology* 7:199-216 (1988).)

15 Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (*J. Biol. Chem* 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over
20 the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (Sec. Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in
25 activity from wild-type.

 Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted
30 form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic

activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions,
5 inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

10 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these
15 positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function.
20 For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are
25 surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions
30 involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glut; replacement of the amide residues Asn and Gln, replacement of

the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention
5 include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for
10 example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of
15 charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev.
20 Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid
25 substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1
30 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is

1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

Polynucleotide and Polypeptide Fragments

5 In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt
10 in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

15 Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-
20 1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini.
25 Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the
30 deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the

invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about"

5 includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the
10 carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are
15 preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-
20 forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide
25 fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable
30 activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these

fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

5 Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second
10 protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous
15 functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the
20 polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-
25 polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the
30

IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion
5 proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified,
10 would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol.
15 Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue,
20 Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767
25 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

30 The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral

vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker
5 for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate
10 promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding
15 portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one
selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin
20 resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as
25 CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-
9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A,
pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and
30 ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1

and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

10 A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most
15 preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical
20 synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also
25 include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some
30 proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids

containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or

translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more
5 restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

10 The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a
15 unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological
20 samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with
25 one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a
30 particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the

present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as
5 molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an
10 immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

15 A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene
20 expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I , ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), and fluorescent labels, such as fluorescein and rhodamine, and
25 biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which
30 emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as

deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I , ^{112}In , $^{99\text{m}}\text{Tc}$), a radio-opaque-substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of $^{99\text{m}}\text{Tc}$. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide.

5 such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from
10 a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

15 The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

20

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune
25 cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide
30 or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic

anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, 5 Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide 10 or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ 15 rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of 20 T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic 25 and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or 30 IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response.

Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Nocardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae,

Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related
5 infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin
10 diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not
15 limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g.,
20 dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present
25 invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against
30 infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable

of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell
5 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

10 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations,
15 polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

20 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers.

25 The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

30 Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a

- candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

- A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

- A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

- A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

- A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

- Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95%

identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of
5 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of
10 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the
15 range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide
20 sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

25 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X
30 in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under
5 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which
10 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide
15 sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of
20 at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule
30 comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X
5 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and
10 determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence
15 selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

20 A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X
25 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological
30 sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted
5 protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1;
10 and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at
15 least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise
20 a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1
25 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the
30 amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of

positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence
5 at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

10 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino
15 acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted
20 protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the
25 amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in
30 the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA

clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of

the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1
 5 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an
 10 isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

15

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector.
 20 Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being
 25 isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

<u>Vector Used to Construct Library</u>		<u>Corresponding Deposited</u>
<u>Plasmid</u>		
30	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
	Zap Express	pBK
	lafmid BA	plafmid BA

pSport1	pSport1
pCMVSPORT 2.0	pCMVSPORT 2.0
pCMVSPORT 3.0	pCMVSPORT 3.0
pCR®2.1	pCR®2.1

- 5 Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are
- 10 commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3
- 15 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the fl origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the fl ori generates sense strand DNA and in the other, antisense.
- 20 Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University,
- 25 NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al.,
- 30 Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P - γ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 μl of reaction

mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific

to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

5 **Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide**

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

10 **Example 3: Tissue Distribution of Polypeptide**

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling
15 system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

20 Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed
25 according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This
30 primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and

hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1

mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The
10 QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

15 The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4,
20 containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

25 In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains:
30 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The

origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

- DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.
- The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

- The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

- Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

- The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

- The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without
5 mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive
10 Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4
15 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using
20 a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the
25 above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

30

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus**Expression System**

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector
5 contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak
10 *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such
15 as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

20 Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP)
25 to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a
30 commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

- 5 The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by
- 10 gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

- Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method
- 15 described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the
- 20 transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

- 25 After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell
- 30 culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then

resuspended in a microcentrifuge tube containing 200 μ l of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

5 To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and
10 cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of 35 S-methionine and 5 μ Ci 35 S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

15 Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

 The polypeptide of the present invention can be expressed in a mammalian
20 cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening
sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient
25 transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV1, HIV1 and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

 Suitable expression vectors for use in practicing the present invention include,
30 for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109),

pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

- 5 Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

- The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing
10 cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., *J. Biol. Chem.* 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., *Biochem. et Biophys. Acta*, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., *Biotechnology* 9:64-68 (1991).) Another useful selection marker
15 is the enzyme glutamine synthase (GS) (Murphy et al., *Biochem J.* 227:277-279 (1991); Bebbington et al., *Bio/Technology* 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the
20 production of proteins.

- Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., *Molecular and Cellular Biology*, 438-447 (March, 1985)) plus a fragment of the
25 CMV-enhancer (Boshart et al., *Cell* 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

- 30 Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be
5 modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

10 The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

15 Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded
20 in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM,
25 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 -
30 PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGC
CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAA
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTACATGCGTGGT
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG
5 ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA
ACCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG
10 GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT
GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCAACGCCT
CCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA
TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
15 GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a
20 polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

25 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell
30 Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in

any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

5 The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as
10 described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

 Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method
15 makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody
20 whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

 It will be appreciated that Fab and F(ab')₂ and other fragments of the
25 antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic
30 chemistry.

 For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced

using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening

10 Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution
15 (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered
20 Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x
25 Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods
30 described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45

minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following
 5 tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells
 10 first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl₂ (anhyd); 0.00130 mg/L CuSO₄·5H₂O; 0.050 mg/L of Fe(NO₃)₃·9H₂O; 0.417 mg/L of FeSO₄·7H₂O;
 15 311.80 mg/L of KCl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄·H₂O; 71.02 mg/L of Na₂HPO₄; 4320 mg/L of ZnSO₄·7H₂O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of
 20 Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0
 25 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H₂O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of
 30 Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319

mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of

5 Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for

10 endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

15 On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the

20 polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

25 **Example 12: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The

30 binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u> <u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	<u>STATs</u>	<u>GAS(elements) or ISRE</u>
	<u>IFN family</u>						
5	IFN-a/B	+	+	-	-	1,2,3	ISRE
	IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
	Il-10	+	?	?	-	1,3	
	<u>gp130 family</u>						
10	IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
	Il-11(Pleiotrophic)	?	+	?	?	1,3	
	OnM(Pleiotrophic)	?	+	+	?	1,3	
	LIF(Pleiotrophic)	?	+	+	?	1,3	
	CNTF(Pleiotrophic)	-/+	+	+	?	1,3	
15	G-CSF(Pleiotrophic)	?	+	?	?	1,3	
	IL-12(Pleiotrophic)	+	-	+	+	1,3	
	<u>g-C family</u>						
	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
20	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
25	<u>gp140 family</u>						
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5	GAS
	GM-CSF (myeloid)	-	-	+	-	5	GAS
30	<u>Growth hormone family</u>						
	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
35	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	1,3	
	CSF-1	?	+	+	-	1,3	GAS (not IRF1)
40							

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

10 5':GCGCCTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATGATTTCCTCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTGGCAAAGCCTAGGC:3' (SEQ ID NO:4)

15 PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

20 5':CTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATGATTTCCTCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCGCCCCTAACTCCGCCCAGTTCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT
25 AGGCTTTTGGCAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter
30 molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the
5 assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and
10 stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the
15 positive control wells.

The above protocol may be used in the generation of both transient, as well as, stable transfected cells, which would be apparent to those of skill in the art.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

20 The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell
25 used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and
30 wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing

10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM
5 KCl, 375 uM Na₂HPO₄·7H₂O, 1 mM MgCl₂, and 675 uM CaCl₂. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400
10 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in
15 the 96-well plate (or 1×10^5 cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant
20 according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes,
25 EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or
30 differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

- 5 The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

- 10 Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

- 15 To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

- 20 PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

- 25 Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

- 30 To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating diseases. For example, inhibitors of NF- κ B could be used to treat those

diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:
5':GCGGCAAGCTTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)
15 Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC
20 ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA
CTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTA
TTCCAGAAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA
GCTT:3' (SEQ ID NO:10)

25 Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP
30 cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly,

the NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with Sall and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25

16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small

molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star
5 black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To
load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate
10 is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension.
15 The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

20 For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4
25 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

30 **Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies.

- 5 In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation
10 of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

- 15 Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

- 20 Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or
25 polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar
30 Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen

Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyn plates (20,000/200ml/well) and cultured overnight in complete medium.

- 5 Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from
- 10 Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on
- 15 ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

- Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described
- 20 here.

- Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and
- 25 PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

- The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂⁺ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride,
- 30 pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the

components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

- 5 Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-
10 POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

- Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound
15 peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

- 20 As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other
25 molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

- Specifically, assay plates are made by coating the wells of a 96-well ELISA
30 plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G

plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1

and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place

of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the

scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1
5 $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day , and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 $\mu\text{g/kg/hour}$, either by 1-
10 4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention
15 are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to
20 modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or
25 microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or
30 poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al.,

Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms)

5 unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage

10 injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

15 Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's

20 solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as

25 phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or

30 arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar

alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an

individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

- 5 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the
10 presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

- 15 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

- 20 Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the
25 media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is
30 produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into

target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be

delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

25

Example 28: Transgenic Animals.

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are

30

used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., *Appl. Microbiol. Biotechnol.* 40:691-698 (1994); Carver et al., *Biotechnology (NY)* 11:1263-1270 (1993); Wright et al., *Biotechnology (NY)* 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci., USA* 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, *Mol Cell. Biol.* 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., *Science* 259:1745 (1993)); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., *Cell* 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," *Intl. Rev. Cytol.* 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., *Nature* 380:64-66 (1996); Wilmut et al., *Nature* 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When

it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the

transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Example 29: Knock-Out Animals.

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., *Nature* 317:230-234 (1985); Thomas & Capecchi, *Cell* 51:503-512 (1987); Thompson et al., *Cell* 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (*e.g.*, see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (*i.e.*,

animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

5 The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or

10 intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and

20 Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For

25 example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological

30 function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

- 5 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both
- 10 incorporated herein by reference in their entireties.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>173</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>April 7, 1998</u>	Accession Number <u>209745</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
EUROPE In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="display: flex; justify-content: space-between; align-items: center;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div> <div style="text-align: center; margin-top: 5px;">Authorized officer</div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="display: flex; justify-content: space-between; align-items: center;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div> <div style="text-align: center; margin-top: 5px;">Authorized officer</div>
--	---

ATCC Deposit No. 209745

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application: to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 208745

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>173</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>April 7, 1998</u>	Accession Number <u>209746</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
EUROPE In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE URNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	
For receiving Office use only <input type="checkbox"/> This sheet was received with the international application Authorized officer	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

ATCC Deposit No. 209746

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 209746

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>173</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>March 20, 1998</u>	Accession Number <u>209683</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
EUROPE In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input style="margin-right: 10px;" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 10px;"></div> <div style="border-top: 1px solid black; padding-top: 5px;">Authorized officer</div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input style="margin-right: 10px;" type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 10px;"></div> <div style="border-top: 1px solid black; padding-top: 5px;">Authorized officer</div>
--	---

Form PCT/RO/134 (July 1992)

ATCC Deposit N . 209683

284

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit N . 209683

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
 - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

5 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

10 4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

20 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

25 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

30 9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence
5 included in ATCC Deposit No:Z;
- (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
- (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- 10 (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in
15 ATCC Deposit No:Z;
- (g) a variant of SEQ ID NO:Y;
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the
20 full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

25

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:

- 30 (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

5 17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

 18. A method of diagnosing a pathological condition or a susceptibility to
10 a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

15

 19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

20 (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

 20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

25 (a) contacting the polypeptide of claim 11 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

 21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

30

22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant;
- 5 (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

<110> Human Genome Sciences, Inc., et al.

<120> 97 Human Secreted Proteins

<130> PZ028PCT

<140> Unassigned

<141> 1999-05-05

<150> 60/085,180

<151> 1997-05-10

<150> 60/085,105

<151> 1997-05-10

<150> 60/085,094

<151> 1997-05-10

<150> 60/085,093

<151> 1997-05-10

<150> 60/085,924

<151> 1997-05-10

<150> 60/085,906

<151> 1997-05-10

<150> 60/085,927

<151> 1997-05-10

<150> 60/085,920

<151> 1997-05-10

<150> 60/085,928

<151> 1997-05-10

<150> 60/085,925

<151> 1997-05-10

<150> 60/085,921

<151> 1997-05-10

<150> 60/085,923

<151> 1997-05-10

<150> 60/085,922

<151> 1997-05-10

<160> 412

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1
 gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg 60
 aattcgaggg tgcaccgtca gtcttcctct tcccccaaa acccaaggac accctcatga 120
 tctcccgagc tcttgaggtc acatgcgtgg tgggtggacgt aagccacgaa gaccctgagg 180
 tcaagttaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg 240
 aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact 300
 ggctgaatgg caaggagtac aagtgcagg tctccaacaa agccctccca acccccatcg 360
 agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc 420
 catcccgga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct 480
 atccaagcga catgcgcgtg gagtgggaga gcaatgggca gccggagaac aactacaaga 540
 ccacgcctcc cgtgctggac tccgacggct ccttcttct ctacagcaag ctcaccgtgg 600
 acaagagcag gtggcagcag gggaacgtct tctcatgtc cgtgatgcat gaggctctgc 660
 acaaccacta cgcgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc 720
 gactctagag gat 733

<210> 2
 <211> 5
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> Site
 <222> (3)
 <223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2
 Trp Ser Xaa Trp Ser
 1 5

<210> 3
 <211> 86
 <212> DNA
 <213> Homo sapiens

<400> 3
 gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aatgatttc 60
 cccgaaatat ctgccatctc aattag 86

<210> 4
 <211> 27
 <212> DNA
 <213> Homo sapiens

<400> 4
 gcggcaagct ttttgcaaag cctaggc 27

<210> 5
 <211> 271
 <212> DNA
 <213> Homo sapiens

<400> 5
 ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60
 aaatatctgc catctcaatt agtcagcaac catagtcctg cccctaactc cgcccatccc 120
 gcccctaact ccgcccagtt ccgcccattc tccgcccctt ggctgactaa ttttttttat 180
 ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240
 ttttgagggc ctaggctttt gcaaaaagct t 271

<210> 6
 <211> 32

<212> DNA
<213> Homo sapiens

<400> 6
gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
<211> 12
<212> DNA
<213> Homo sapiens

<400> 8
ggggactttc cc 12

<210> 9
<211> 73
<212> DNA
<213> Homo sapiens

<400> 9
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60
ccatctcaat tag 73

<210> 10
<211> 256
<212> DNA
<213> Homo sapiens

<400> 10
ctcgagggga ctttcccgga gactttccgg ggactttccg ggactttcca tctgccatct 60
caattagtag gcaaccatag tcccgcccct aactccgccc atcccgcccc taactccgcc 120
cagttccgcc cattctccgc cccatggctg actaattttt ttattttatg cagaggccga 180
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg 240
cttttgcaaa aagctt 256

<210> 11
<211> 1191
<212> DNA
<213> Homo sapiens

<400> 11
gctgggctgg aacacaagar cccacagggc tgccgtccac actctcccg tccagatcct 60
gggaccacat ggggacgctg ccatggcttc ttgccttctt cattctgggt ctccaggctt 120
gggatactcc caccatcgtc tcccgcaagg agtggggggc aagaccgctc gcctgcaggg 180
ccctgctgac cctgctgtg gcctacatca tcacagacca gctcccaggg atgcagtgc 240
agcagcagag cgtttgcagc cagatgctgc gggggttgca gtccattcc gtctacacca 300
taggctgggtg cgacgtggcg tacaacttcc tgggtgggga tgatggcagg gtgtatgaag 360
tggttggctg gaacatccaa ggcttgaca cccagggcta caacaacatt tccctgggca 420
tcgcttctt tggcaataag ataagcagca gtccagccc tgctgcctta tcagctgcag 480
agggctctgat ctctatgcc atccagaagg gtcacctgc gccaggtat attcagccac 540
ttctctgaa agaagagacc tgctggacc ctcaacatcc agtgatgcc agraaggtt 600

```

gccccaaacat catcaaacga tctgcttggg aagccagaga gacacactgc cctaaaatga 660
acctcccagc caaatatgtc atcatcatcc acaccgctgg cacaagctgc actgtatcca 720
cagactgccca gactgtcgtc cgaacacatac agtcctttca catggacaca cggaactttt 780
gtgacattgg atatcaataa ggccaggcgt ggccgctgatt acgtctgtaa tcccaggact 840
ttggggaggcc aaggcgggca gatcacttca ggccaggaat tcaagagcag cctggccaat 900
atggcgaaac tctgtctcta ctgaaaacaa acaaacaaac aaacaaacaa acaaagaaac 960
aacaataatt agccgggtgt ggtggcacac gcctgtagtc ccagctactc aggaggctga 1020
ggcataagaa ttgcttgaac cctggaggcg gaggttgca ttagctgaga ttggggccacc 1080
gcactccagt ctggggagaca gagtgcagct gtctcaaac aacaacaaaa aaatccctaa 1140
cataatctca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa agggcggccg c 1191

```

<210> 12
 <211> 1251
 <212> DNA
 <213> Homo sapiens

```

<400> 12
ggcagaggtc agccaactaa caaatgaagc gcagggaat gactcaattc ttattgagtc 60
tagttgctct taattgctgc tctatttctt tgggaagatt gacatatcca ggagggtttc 120
atctaaaact agacccttta gaactctgaa gtcagagcaa ctttccctct gtcaatccta 180
ctcactactt ttgtamcctt gaccagagaa gttgcttaat cttttggggc ctgcattctc 240
atatacctaa agtaggaata aaaatacctg cttagagact tgctcagtc atcaaatrag 300
agattatata caaccttccc acttcaagga tggctgcaag gacaaaaaag aaaaatgaca 360
taataaataa aaaggtccct gcagactgta atactaggat gaggttattac tacaagggtc 420
caggggaaaag aggagagatg gagtcttggt tggcatgct atcatggtct attttagatt 480
ttgagttttt agaggcaaga ccacagttgt ttaatttagt gtatacagaa cattccactt 540
attcagggag acattatact agggaaaggg gtgggttcat ggtgttcaaa aattcact 600
cacagttatt attaaaaaga aaggattctc tatgtgcttt tttcagccc atggctttaa 660
atatcatcca tgtgcctatg tcttccaaat gtatttttcc agcccagctc ggtccctcga 720
cattcagatc cttatgggtg tgccctcacc ctatatccaa atggcaactt ggtctctact 780
ctagtcagat tagagatata ccatacttgg catgactaaa atggaaactt aacttgtttc 840
ttatctctat ctacgtaaat cacaccacca cagtgcatac ttttctctaa tcaaattcct 900
aagaatcatc cttgattttt cccttctctt tgtcccttgc catcccagat tatcctgcaa 960
aaactgtcta tgctacctac aaaagtatct gccacatgct atactaattg tcratatcct 1020
agagcaccmt tcatctgcct tcacctgtgg tgttgctgca attgtctcct tectggctgc 1080
cctgattata tccattctcc ctgtctccaa aagcattctg cgcacagcag acacagatgt 1140
ttcataaatg taagtctggt catgcgctcc tctacctaaa accattagat ggtttttcat 1200
tgcactcaca actagagttt cctgaccatg acttgaggc taagctcgta g 1251

```

<210> 13
 <211> 1734
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1417)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1703)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1714)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1715)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1731)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1732)

<223> n equals a,t,g, or c

<400> 13

gaagcgtg	gtgccgcg	aatggcgg	ctcacatt	ccacgggt	tggcaatt	60
ttttcggt	tggcgctc	ggtgactc	ctcaaatg	ttctcatc	cacatacc	120
tccacagat	ttgaagtac	ccgaaact	cttgctat	ctcacagt	gccaatat	180
cagtggtat	atgaggca	ttcagagt	acgttggt	acccccct	ctttgcat	240
tttgagtata	tcctgtcac	tgttgccaa	tattttgat	aagaaatg	gaatgtcc	300
aatttgaat	actccagtc	aaggacct	ctttccaga	gattttcc	catctttat	360
gatgtactc	ttgtgtat	tgctcggt	tgctgtaa	gcattgat	aaaaaaag	420
ggtaaaga	ttacagaaa	gccaaaatt	attctgtcg	tattacttc	gtggaact	480
gggttatt	ttgtggacc	tattcatt	cagtacaat	gctttttat	tggattaat	540
ctactctcc	ttgcacgat	atttcagaa	aggcatat	aaggagcat	tctctttgt	600
gttctcct	atttcaag	tatctacct	tatgtagca	cagcttat	tgtatatct	660
ctgcgatct	actgtttca	tgcaataaa	ccagatgg	ctattcgat	gaagagtt	720
agctttgt	gtgttatt	cctgggact	gttgtttct	tagtttct	tctttcatt	780
ggctcttcc	tggccttga	tcagctgc	caagtcttt	cccgaact	tcctttca	840
aggggcct	gtcatgcata	ttgggctcc	aacttctgg	ctttgtaca	tgctttgg	900
aaagtgtgt	ctgtcatcg	tttgaatt	aaatttct	atcccaaca	tattccca	960
gcctcaat	caagtgtgt	ggttcagca	ttccaacac	cagtccttc	ctcagtgat	1020
cccttggca	ccctcatct	cacactgat	gccatatt	cctctatt	ctgtcttt	1080
tttaaaccc	aagggcccg	aggctttct	cgtatgtct	ctctttgt	cttgagctc	1140
tttatgttg	ggtggcatg	tcatgaaaa	gccatact	tagcaatt	cccaatgag	1200
cttttgtct	tgggaaaag	aggagacgt	tcgatttt	tgattctga	cacaacagg	1260
cattattccc	tcttccct	gctcttcat	gcaccaga	ttccattaa	aatcttact	1320
atgttactat	tcaccatata	tagtatttc	tcactgaag	ctttattcag	aaaagaaaa	1380
cctcttttta	attggatgg	aaactttct	ctgcttngc	tggggcctc	ggaagctgc	1440
tgtgaatttg	tattccctt	caactcctg	aagggtga	acccctcat	ccctttgtt	1500
ctaacctcag	tgtattgtc	agtaggcac	acatatgct	ggttcaaat	gtatgtttc	1560
gtattgattg	actctgctat	tggcaagaca	aagaaaca	gaataaagg	actgcttag	1620
aaaaaaaaa	aaaaaaaaa	aaagggcgg	cgctctag	gatccctcg	gggcccaag	1680
ttacgcgtg	atgcgagtc	tantctctc	tggntgat	gtatgaagc	nhgc	1734

<210> 14

<211> 1540

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (22)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (430)

<223> n equals a,t,g, or c

<400> 14

gcctgggcg	cgtgggcg	gnactgcg	ggctgcg	gtgccgag	gcgcgagg	60
cggggaagg	gcacctggg	tggccctgg	gtgcggcg	cgacatgg	gacggcgtg	120
tcaaggagg	cttctgggt	aagaggggc	acattgtcc	caactgga	gcgcgatgt	180
tcacctctg	gcagaacac	ctggtgtat	acaagcttg	ggggggtcg	agagtgaac	240

ctcccaaggg	cggatcctc	ctggatgggt	gcacntcac	ctgcccctgc	ctggagtatg	300
aaaaccgacc	gctcctcatt	aagctgaaga	ctcaaacatc	cacggagtac	ttcctggagg	360
cctgttctcg	agaggaagcg	ggatgcctgg	gcctttkaag	rt yaccggg	ctattcatgc	420
agggcagccn	ggggaagggtc	cagcagctgc	acagcctgag	aaactccttc	amgctgcccc	480
cgcacatcar	gctgyatcgy	attgtggaca	agatgcacga	tagcaacacc	ggwatccggt	540
caagccccc	catggagcag	agaagcacct	ataaaaagam	cttyctcggc	tcctccctgg	600
tggactggyt	yatctycaam	agcttcamgg	gcagccgtct	kgaggcgggtg	amcctggcct	660
cgatgytcat	rgaggagaac	ttcctcaggt	ctgtggctgt	acgatgcattg	ggaggcattc	720
ggctctgggga	tctggccgag	cagttcctgg	atgactccac	agccctgtac	acttttsctg	780
agagctacam	aaagawgata	agccccaagg	aagaaattag	cctgagcact	gtggagttaa	840
gtggcagcgt	ggtgaaacaa	ggctacctgg	ccaagcaggg	acacaagagg	aaaaactgga	900
cgttgcgtcg	ctttgttcta	aggaaggatc	cagctttcct	gcattactat	gacccttcca	960
aagaagagaa	caggccagtg	ggtgggtttt	ctcttcgtgg	ttcactcgtg	tctgctctgg	1020
aagataatgg	cgttcccact	ggggttaaag	ggaatgtcca	gggaaacctc	ttcaaagtga	1080
ttactaagga	tgacacacac	tattacattc	aggccagcag	caaggctgag	cgagccgagt	1140
ggatggaagc	tatcaaaaag	ctaacaatgac	aaggacctga	gggaaccagg	attcctccct	1200
cctaccagat	gacacagaca	agagtctctg	gagaatggga	gtgttaagac	tttgacttc	1260
tttgaagtgt	ttgtactgct	ttggagagtg	aatgtctgcca	agagtctctc	agattacaaa	1320
cagcagtggt	gccatttctt	tcccattctt	catgttacaa	acctggaaaag	gctagaacag	1380
ccattagcgt	tcagcatctt	gacttttccc	cagcatcaca	aacagccatt	tcctcggcca	1440
ccaaagtagg	ttccctttgt	tggaaacaatt	acactggcca	tgccataatg	ttgaataaaa	1500
ctctcttctt	atgaaaaaaa	aaaaaaaaaa	aaaaaaaaaa			1540

<210> 15

<211> 1558

<212> DNA

<213> Homo sapiens

<400> 15

ccacgtcgtc	cgaacctttt	aaaaatgggtc	ttgatgtatg	tggaaagagag	tatgtgtatg	60
tgtgttccctg	tacatagcat	gggtgcagct	gtggatgtgt	gcaaaagagt	gtgagtgtgt	120
gtgtgtgtgt	gtaaaggggt	ctgtccctaga	gcccacatca	gtttgttgtg	aatctggaaa	180
aagggtcgggt	gagggccggg	agatgttgac	cctggtggga	gcaggctgag	gctgccccgt	240
tctccacatc	ctctgttttg	cccagttctt	gattccatta	gggggagtg	gctgaagcca	300
ttctcggatg	cttcccagac	caggctccct	ctgccagagt	cacatgcac	cgagctgctg	360
gtctccattg	tccagcagga	aggcggaaag	gcaggcaaga	tgggtgtgaag	cttaaagctt	420
gtatttgatg	gaaaagggtct	cccctgttca	tctgagaggc	caagcctggc	caccccaggc	480
tcagaacctg	ggcttcaaga	aatgtgctgg	gagctcctaa	cttacacatc	cctccagcct	540
tccttgaaac	ctcccaccac	cccctatttc	ctttaatttc	tcagggtctg	tcctctctcc	600
cccaacccca	cagctgggca	agaagtctgc	aaaagctgca	tctgcagctg	tctctaactc	660
ttcccagcca	tctcccgtat	tttttggtac	cttgattcct	tgactcttaa	taagccaagc	720
caccttatct	ctgtagtctt	tatttttttg	ttgactaaat	ttgggggggt	cttttttatg	780
gtcatgtcac	tgacctatta	aattgggggt	tggtgctttt	ccaccttccc	cctctgaatg	840
aaagccaagg	aatgggggaa	gagcgggaac	tctgccgcgg	aggtggagca	agaacgggtga	900
agggccctgg	tcccagagag	gctggtgggt	ccctctccca	aaggaaggca	gacagtctct	960
gctttgcctt	ggaccttggt	gctgggggtg	gggaggcctg	ggggggacac	tcctccactcc	1020
cattcccctt	cctttgtcct	aatcctggaa	ttaagtacag	gggtttatag	gttctatttc	1080
ttcccaagag	ccctgcaaag	aaccccagtt	tcctatttgg	atgcccctac	actgttgtgt	1140
ttcagtggaa	tgtattttca	tttaaaaaaca	actttgaatg	gggcactttt	tctttccctg	1200
tttaaaaatt	gaaaaattct	tacagtacaa	acaggactgt	cagggtgggg	gtgttggtgc	1260
tgtaagaggt	tactcttgag	tgcattttgg	cactgggatg	ggatggctgg	ggtgggaaga	1320
cccccatccc	cacccccaac	ttcttttcta	atatttaagy	agtgttttgt	aggattcaac	1380
aaccaccaca	acttgaattt	gtatcatggg	aggtgggagg	gagtggctta	gaggtgtctg	1440
cctatgctta	aagccaactg	tggaaagttt	gttttccctt	ttttgtataa	taaagtgaag	1500
aacaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1558

<210> 16

<211> 1636

<212> DNA

<213> Homo sapiens

<220>

<221> SITE
<222> (424)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (823)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (960)
<223> n equals a,t,g, or c

<400> 16
gaattcggca cgagttgaaa ttgaaaatca agataaaaat gttcacaatt aagctccttc 60
tttttattgt tcctctagtt atttcctcca gaattgatca agacaattca tcatttgatt 120
ctctatctcc agagccaaaa tcaagatttg ctatgttaga cgatgtaaaa attttagcca 180
atggcctcct tcagttggga catgggtcta aagactttgt ccataagacg aagggccaaa 240
ttaatgacat atttcaaaaa ctcaacatat ttgatcagtc tttttatgat ctatcgctgc 300
aaaccagtgga aatcaaaagag gaagaaaagg aactgagaag aactacmtat aaactacaag 360
tcaaaaatga agaggtaaag aatatgtcac ttgaactcaa ctcaaaactt gaaagcctcc 420
tagnagaaaa aattctactt caacaaaaag tgaaatatat agaagagcaa ctaactaact 480
taattcaaaa tcaacctgaa actccagAAC acccagaagt aacttcactt aaaacttttg 540
tagaaaaaca agataatagc atcaaaagacy ttctccagac cgtggaagac caatatwaac 600
aattaaacca acagcatagt caaataaaag aratagaaaa tcagctcaga aggactagta 660
ttcaagaacc cacagaaatt tctctatctt ccaagccaag agcaccaaga actactccct 720
ttcttcagtt gaatgaaata agaaatgtaa aacatgatgg cattcctgct gaatgtacca 780
ccattttataa cagaggtgaa catacaagtg gcatgtatgc atncagaccc agcaactctc 840
aagtttttca tgtctactgt gatgttatat caggtagtcc atggacatta attcaacatc 900
gaatagatgg atcacaaaac ttcaatgaaa cgtgggagaa ctacaaatat ggttttggn 960
aggcttgatg gagaattttg gttgggccta gagaagatat actccatagt gaagcaatct 1020
aattatgttt tacgaattga gttggaagac tggaaagaca acaaacatta tattgaatat 1080
tctttttact tgggaaatca cgaaccaaac tatacgtac atctagttgc gattactggc 1140
aatgtcccca atgcaatccc ggaaaacaaa gattttggtgt tttctacttg ggatcacaaa 1200
gcaaaaggac acttcaactg tccagaggggt tattcaggag gctggtggtg gcatgatgag 1260
tgtggagaaa acaacctaaa tggtaaatat aacaaaccaa gagcaaaatc taagccagag 1320
aggagaagag gattatcttg gaagtctcaa aatggaaggt tatactctat aaaatcaacc 1380
aaaaatgttg tccatccaac agattcagaa agctttgaat gaactgaggc aaatttaaaa 1440
ggcaataatt taacatttaa cctcattcca agttaatgtg gtctaataat ctggatttaa 1500
atccttaaga gaaagcttga gaaatagatt ttttttatct taaagtcact gtctatttaa 1560
gattaaacat acaatcacat aaccttaaaa aaaaaaaa aaaaactcga ggggggcccc 1620
gtaccaatc cgccgg 1636

<210> 17
<211> 1256
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1240)
<223> n equals a,t,g, or c

<400> 17
tcgaccacag cgtccgagca accgcagctt ctagtatcca gactccagcg ccgccccggg 60
cgcggaaccc aaccccgacc cagagctctt ccagcgggcg cgacagagca gggctccccg 120
ccttaacttc ctccgcgggg ccagccacc ttcgggagtc cgggttgccc acctgcaaac 180
tctccgctt ctgcacctgc cacccttgag ccagcgggcg cgcccgagcg agtcatggcc 240
aacgcggggc tgcagctgtt gggcttcatt ctgccttcc tgggatggat cggcgccatc 300
gtcagcactg cctgcccga gtggaggatt tactcctatg ccggcgacaa catcgtgacc 360
gccaggcca tgtacgagg gctgtggatg tctgctgtgt cgacagacac cgggcagatc 420
cagtgcaaaag tctttgactc ctgtctgaat ctgagcagca cattgcaagc aaccctgtgc 480

ttgatggtgg	ttggcctcct	cctgggagtg	atagcaatct	ttgtggccam	cgttggcatg	540
aagtgtatga	agtgccttga	agacgatgag	gtgcagaaga	tgaggatggc	tgatcattggg	600
ggcgcgatat	ttcttcttgc	aggtctggct	attttagttg	ccacagcatg	gtatggcaat	660
agaatcggtc	aagaattcta	tgacctatg	acccagtc	atgccaggta	cgaatttggg	720
caggctctct	tcactggctg	ggctgctgct	tctctctgcc	ttctgggagg	tgccctactt	780
tgctgttctc	gtccccgaaa	aacaacctct	tacccaacac	caaggcccta	tccaaaacct	840
gcaccttcca	gcgggaaaga	ctacgtgtga	cacagaggca	aaaggagaaa	atcatgttga	900
aacaaaccga	aaatggacat	tgagatacta	tcattaacat	taggacctta	gaattttggg	960
tattgtaatc	tgaagtatgg	tattacaaaa	caaacaacaa	aacaaaaaac	ccatgtgtta	1020
aaatactcag	tgctaaacat	ggcttaatct	tattttatct	tctttctcca	atataggagg	1080
gaagattttt	ccattttgat	tactgcttcc	cattgagtaa	tcatactcaa	ctgggggaag	1140
gggtgctcct	taaataatata	tagatatgta	tatatcatg	tttttctatt	aaaaatagac	1200
agtataaatwc	taaaaaaaaa	aaaaaaaaamcy	cggggggggg	ccggtaccca	ttcgcc	1256

<210> 18

<211> 1143

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1100)

<223> n equals a,t,g, or c

<400> 18

ggcacgaggg	ctgggggtcag	caaataatca	gggggcccag	gcgtcacgtg	ggccccatcc	60
tcagcagcag	tgccctcgat	atcttctgcg	acaatgagaa	tgggccctaac	ttccttttcc	120
acaaccgggg	cgatggcacc	tttgtggacg	ctgcggccag	tgctggtgtg	gacgaccccc	180
accagcatgg	gcgaggtgtc	gccctggctg	acttcaaccg	tgatggcaaa	gtggacatcg	240
tctatggcaa	ctgggaatggc	ccccaccgcc	tctatctgca	gatgagcacc	catgggaagg	300
tccgcttccg	gggacatcgc	cttcacccaa	gttctccatg	ccctcccctg	ttccgacagg	360
tcataccggg	ccgactttga	caatgaccag	gagctggaga	atcttcttca	acaacattgc	420
ctaccgcagc	tcctcagcca	accgcctctt	ccgcgtcatc	cgtagagagc	acggagaccc	480
cctcatcgag	gagctcaatc	ccggcgacgc	cttggagcct	gagggccggg	gcacaggggg	540
tgtggtgacc	gacttcgacg	gagacgggat	gctggacctc	atcttgtccc	atggagagtc	600
catggctcaa	ccgctgtccg	tcttccgggg	caatcagggc	ttcaacaaca	actggctgcg	660
agtgggtgca	cgcacccggg	ttggggcctt	tgccaggggg	gctaaggctc	tgctctacac	720
caagaagagt	ggggcccacc	tgaggatcat	cgacgggggc	tcaggctacc	tggtgtgagat	780
ggagcccgtg	gcacactttg	gcctggggaa	ggatgaagcc	agcagtggtg	aggtgacgtg	840
gccagatggc	aagatgggtga	gccgggaacg	ggccagcggg	gagatgaact	cagtgtctga	900
gatcctctac	ccccgggatg	aggacacact	tcaggaccca	gccccactgg	agtgtggcca	960
aggattctcc	cagcagggaa	atggccattg	catggacacc	aatgaatgca	tccagttccc	1020
attcgtgtgc	cctcgagaca	agcccgtatg	tgtcaacacc	tatggaagct	acaggtgccg	1080
gaccaacaag	aagtgcagtn	cgggggtacg	agtcccaacg	aggatggcac	atacgggctt	1140
gtc						1143

<210> 19

<211> 1537

<212> DNA

<213> Homo sapiens

<400> 19

atcatatagg	aaacggtagc	ctgcagtacc	gggtccggaat	ttccgggtcg	acccacgcgt	60
ccggagcagc	aagagatttg	tcctggggat	ccagaaaccc	atgataccct	actgaacacc	120
gaatcccctg	gaagcccaca	gagacagaga	cagcaagaga	agcagagata	aatacactca	180
cgccaggagc	tcgctcgctc	tctctctctc	tctctcactc	ctccctccct	ctctctctgc	240
ctgtcctagt	cctctagtcc	tcaaattccc	agtcctctgc	accccttcc	gggacactat	300
gttgttctcc	gcctcctctc	tgagagtgat	ttggatcctg	gctgcagatg	ggggtcaaca	360
ctggacgtat	gagggcccac	atggtcagga	ccattggcca	gcctcttacc	ctgagtgtgg	420
aaacaatgcc	cagtcgcccc	tcgatattca	gacagacagt	gtgacatttg	accctgattt	480
gcctgctctg	cagccccacg	gatatgacca	gcctggcacc	gagccttttg	acctgcacaa	540
caatggccac	acagtgaac	tctctctgcc	ctctaccctg	tatctgggtg	gacttccccg	600

aaaatatgta gctgcccagc tccacctgca ctggggctcag aaaggatccc caggggggtc	660
agaacaccag atcaacagtg aagccacatt tgcagagctc cacattgtac attatgactc	720
tgattcctat gacagcttga gtgaggtctc tgagaggcct cagggcctgg ctgtcctggg	780
catcctaatt gagctggaaa agcttcaggg gacattgttc tccacagaag aggagccctc	840
taagcttctg gtacagaact accgagccct tcagcctctc aatcagcgca tggcttttgc	900
ttctttcatc caagcaggat cctcgtatac cacagggtgaa atgctgagtc taggtgtagg	960
aatcttggtt ggctgtctct gccttctcct ggctgtttat ttcattgcta gaaagattcg	1020
gaagaagagg ctggaaaacc gaaagagtgt ggtcttcacc tcagcacaag ccacgactga	1080
ggcataaatt ccttctcaga taccatggat gtggatgact tcccttcattg cctatcagga	1140
agcctctaaa atgggggtgta ggatctggcc agaaacactg taggagtagt aagcaatgt	1200
cctccttccc ctggacatct cctagagagg aatggaccca ggctgtcatt ccaggaagaa	1260
ctgcagagcc tttagcctct ccaaactatg aggaggaaat gaggaaatcg ctgtgttgtt	1320
aatgcagaga acaaactctg tttagttgca ggggaagtgt ggatataacc ccaaagtctt	1380
ctacccctc acttttatgg ccttttccct agatatactg cgggatctct ccttaggata	1440
aagagttgct gttgaagttg tatatttttg atcaatatat ttggaaatta aagtttctga	1500
cttataaaaa aaaaaaaaaa aaaaaactcg agggggg	1537

<210> 20
<211> 2672
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (16)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (28)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (47)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (52)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (93)
<223> n equals a,t,g, or c

<400> 20	
cccaaagttc ggaaantaaa ccttcaanta aagggaaaca aaaagcngga gnttcccacc	60
gcgggtgggc ggcccgttct agaattaagt ggnatcccc cggggctgcc aggaatttcc	120
gagccggggc cgcgccgccg ctgcccgccg ccgcgsgcgg attytgcttc tcagaagatg	180
cactattata gatactctaa cgccaaggtc agctgctggt acaagtacct ccttttcagc	240
tacaacatca tcttctgrtt ggctggagtt gtcttccttg gagtcgggct gtgggcatgg	300
agcgaaaagg gtgtgctgtc cgacctcacc aaagtgacct ggatgcatgg aatcgacctc	360
gtgtgctggt tcctgatggt gggcgtgggt atgttcaccc tggggttcgc cggctgcgtg	420
ggggctctgc gggagaatat ctgcttgctc aactttttct gtggcaccat cgtgctcctc	480
ttcttctggt agctggctgt ggccgtgctg gccttccctg tccaggactg ggtgagggac	540
cgggtccggg agttcttctga gagcaacatc aagtcctacc gggacgatat cgtctgcaa	600
aacctcatcg actcccttca gaaagctaac cagtgtgtg gcgcataatg cctgaagac	660
tgggacctca acgtctactt caattgcagc ggtgccagct acagccgaga gaagtgcggg	720
gtcccttctt cctgctgctg gccagatcct gcgcaaaaag ttgtgaacac acagtgtgga	780
tatgatgtca ggattcagct gaagagcaag tgggatgagt ccatcttcac gaaaggctgc	840
atccaggcgc tggaaagctg gctcccgccg aacatttaca ttgtggctgg cgtcttcctc	900

gcatctctgc	tgttgcagat	atttggcatc	ttcttggcaa	ggacgctgat	ctcagacatc	960
gaggcagtg	aggccggcca	tcacttctga	ggagcagagt	tgagggagcc	gagctgagcc	1020
acgctgggag	gccagagcct	ttctctgcca	tcagccctac	gtccagaggg	agaggagccg	1080
acaccccccag	agccagtgcc	ccatcttaag	catcagcggtg	acgtgacctc	tctgtttctg	1140
cttctgtggtg	ctgaagacca	aggggtcccc	ttgtttacctg	cccaaacttg	tgactgcac	1200
cctctggagt	ctaccagag	acagagaatg	tgtctttatg	tgggagtggt	gactctgaaa	1260
gacagagagg	gtctctgtgg	ctgccaggag	ggcttgactc	agacccctg	cagctcaagc	1320
atgtctgcag	gacacccctg	tcccctctcc	actggcatcc	agacatctgc	tttgggtcat	1380
ccacatctgt	gggtgggccc	tgggttagagg	gacccacagg	cgtggacagg	gcactctctc	1440
ccatcaagca	aagcagcatg	ggggc-tgcc	cgtaacggga	ggcggacgtg	gccccgctgg	1500
gcctctgagt	gccagcgcag	tctgtctggga	catgcacata	tcagggggtg	tttgcaggat	1560
cctcctcag	gttcaagtga	agtaagcctg	agccagtgcg	tggactgggtg	ccacgggagt	1620
gccttgtcca	ctgtccccct	gtgtccacca	gctattctcc	tggcgccgga	actgcctctg	1680
gtcttgatag	cattaagccc	tgatggcgcc	gggtggcggtt	gggcatgggtt	cttctactgag	1740
agccggctct	ccttttctta	aagtgtgtaa	atagtttatt	tataggggta	agaatgttct	1800
cacacattt	cacttctctc	tctctctctc	cagcattctc	ctctgagcag	ccttagatag	1860
tgtccatggc	tggagccgac	cctttgagtc	cccttgagt	tcttaagaac	cagcccaaaa	1920
cagcctctct	ttctctctca	catactgcag	cctccctcca	tgcacccac	atacaagcac	1980
tccccactc	cccagcgtgg	cctcactgtc	ttctgggtctt	gggtgctactg	aaattgtcac	2040
ccagaatttg	aatcctgacc	ctccccactg	caagcccagg	gagccccagc	ccaagatggc	2100
cagcctgaaa	ctgttggcca	gggtcctctc	tgtggccatg	tacccagggc	tggctggcct	2160
gccatttggc	tctccccgga	gacagccgtt	cttctgcaac	cacacccctg	gcctagccac	2220
aacccccagg	tgcagctgct	cagaagctcc	aggcattttg	tttctgggtga	ccgccccata	2280
tgggatattg	gtgatcactg	gtccaccctt	cctgtcaggg	ctttctctgg	gctgctcttg	2340
gaatgaagt	cttaagtact	gaataactcc	cctggggata	gctggggcat	ttgtctagct	2400
gggtactttt	ctaactttt	gccatagctc	agaccacttc	tcacgttca	gggatggact	2460
gcaaccttaa	tttacttgcc	ggagtgtaca	ttctagtgtg	gtgtatactg	gtggctgttg	2520
atgatgattt	ttttttttt	tttacacaat	tctctgtaga	ctaggagaag	aatgcttgtg	2580
tttttcggaa	gtgtgatgct	tctctttgac	tgccaaactc	ttttatggaa	tatatcttta	2640
tattaaaaaa	aaaaaaaaac	aaaaaaaaaa	aa			2672

<210> 21

<211> 1508

<212> DNA

<213> Homo sapiens

<400> 21

ggcacagaga	tagagcggca	acctcggaag	tgcggacggg	tgggcctata	tagatgttga	60
gggtgcggagg	ccgtgggctt	ttgttgggccc	tggctgtgac	cgagcagcg	gtaatggcag	120
cacggccttat	gggtctgttg	gggtccccgcg	ctggctttcg	ccttttcata	ccggaggagc	180
tgtctcgtcta	ccgcggcggc	ccaggggacc	cgggcctgta	cttggcgctg	ctcgccctg	240
tctacgatgt	gtcctccggc	cggagcacta	cgagcctggg	tccactata	gcggcttcgc	300
aggccgagag	gcattccagag	ctttcgtgac	cggggactgt	tctgaagcag	gcctcgtgga	360
tgacgtatcc	gacctgtcag	ccgctgagat	gctgacactt	cacaattggc	tttctattcta	420
tgagaagaat	tatgtgtgtg	ttgggagggt	gacaggacgg	ttctacggag	aggatgggct	480
gccccacccg	gcactgaccc	aggtagaagc	tgcgatcacc	agaggtcttg	aggccaacaa	540
actacagctg	caagagaagc	agacattccc	gccgtgcaac	gcggagtggg	gctcagccag	600
gggcagccgg	ctctggtgct	cccagaagag	tggagggtgtg	agcagagact	ggattggcgt	660
ccccagggaag	ctgtataagc	caggtgctaa	ggagccccgc	tgcgtgtgtg	tgagaaccac	720
cggccccctt	agtggccaga	tgccggacaa	ccctccacac	agaaatcgtg	gggacctgga	780
ccacccaaac	ttggcagagt	acacaggctg	cccaccgcta	gcatcactat	gctcctttcc	840
actctaagcc	gtagcctctt	ctgttaataa	cacacagaga	gctctgcca	gcacctgagt	900
aggcccttga	cacttgtgtg	ccctgggatg	cctcctggcg	cgaatcagga	gggtctggaa	960
ggactctggc	tatatctctg	aaatgtggct	catgcccctt	acgctggctc	ggcgttgtgtg	1020
tgcctgaggg	acagccggcc	acctgcccag	tactgtgtcag	cttttcaaca	ctattccctt	1080
tgacctactg	gccattcttc	tcacagccct	cagatatcaa	cgggcacaaa	taagaccaac	1140
tcaatttcca	cttgaattta	caacccaaaag	cctgctgagt	tgattacagc	tgggccaata	1200
cagtacgagg	caataacaaa	ttagtgtggg	ttgattcttg	aattggaaaa	gcttttgcct	1260
gtatggatac	agcaaatcca	gatgtctctg	aacaaaagca	caattttaaag	caacgacatt	1320
ttctgtcctt	taagcactta	aaatcagggtg	tgggtgtgtt	tcaaaggcag	aagtctgcat	1380
tttgagcaaa	aggtggcttc	ccagctctaa	caaggtaact	ggttagcatg	acattaaagc	1440
ttgggcaagg	cttcaaaactt	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1500
aactcgag						1508

<210> 22
<211> 1447
<212> DNA
<213> Homo sapiens

<400> 22
aattcggcac gagagattta agtgcagcgt ggattttttt tttctcactt tgccttgtgt 60
tttccaccct gaaagaatgt tgtggctgct cttttttctg gtgactgccca ttcattgctga 120
actctgtcaa ccagggtgcag aaaatgcttt taaagtgaga cttagtatca gaacagctct 180
gggagataaa gcatatgcct gggataccaa tgaagaatac ctcttcaaag cgtatggtagc 240
tttctccatg agaaaagttc ccaacagaga agcaacagaa atttcccatg tctactttg 300
caatgtaacc cagaggggtat cattctgggt tgtggttaca gacccttcaa aaaatcacac 360
ccttctgct gttgaggtgc aatcagccat aagaatgaac aagaaccgga tcaacaatgc 420
cttctttcta aatgmccaaa ctctggaatt tttaaaaatc ccttccacac ttgcaccacc 480
catggacca tctgtgcccc tctggattat tatatttgggt gtgatatttt gcatcatcat 540
agttgcaatt gcactactga ttttatcagg gatctggcaa cgtagaagaa agaacaaaga 600
accatctgaa gtggatgacg ctgaagataa gtgtgaaaac atgatcacaa ttgaaaatgg 660
gatccctct gatccctctg acatgaaggg agggcatatt aatgatgcct tcatgacac 720
ggatgagagg ctcaccctct tctgaagggc tgttgttctg ctctctcaag aaattaaaca 780
tttgtttctg tgtgactgct gagcatcctg aaataccaag agcagatcat atattttgtt 840
tcaccattct tcttttgtaa taaattttga atgtgcttga aagtgaagaa caatcaatta 900
taccaccaa caccactgaa atcataagct attcacgact caaaatattc taaaatattt 960
ttctgacagt atagtgtata aatgtgttca tgtgttattt gtagttattg atttaagcat 1020
ttttagaaat aagatcaggc atattgtatat attttcacac ttcaaagacc taaggaaaaa 1080
taaattttcc agtggagaat acatataata tgggttagaa atcattgaaa atggatcctt 1140
tttgacgac acttatatca ctctgkatat gactaagtaa acaaaagtga gaagtaatta 1200
ttgtaaatgg atggataaaa atggaattac tcatatacag ggtggaattt tatctgttta 1260
tcacaccaac agttgattat atattttctg aatatcagcc cctaatagga caattctatt 1320
tgttgacctt ttctacaatt tgtaaaagtc caatctgtgc taacttaata aagtaataat 1380
catctctttt tgatttgtgaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1440
actcgag 1447

<210> 23
<211> 1583
<212> DNA
<213> Homo sapiens

<400> 23
ggcagcagg acaacgacta tctgtctacat ggtcatagac ctcccatgtt ctcttttcgg 60
gcttgcttca agagcatctt ccgcattcat acagaaactg gcaacatctg gaccatctg 120
cttggtttct tgctgtttct ctttttggga atcttgacca tgcctcagacc aaatatgtac 180
ttcatggccc ctctacagga gaaggtggtt tttgggatgt tctttttggg tgcagtgtc 240
tgcctcagct tctctggct ctttcacacc gtctattgtc attcagagaa agtctctcgg 300
actttttcca aactggacta ttcagggatt gctcttctaa ttatggggag ctttgtcccc 360
tggctctatt attccttcta ctgctcccca cagccacggc tcatctacct ctccatcgct 420
tgtgtcctgg gcattttctg catcattgtg gcgcagtggt accggtttgc cactcctaag 480
caccggcaga caagagcagg cgtgttctct ggacttggct tgagtggcgt cgtgccacc 540
atgcacttta ctatcgctga gggctttgtc aaggccacca cagtgggcca gatgggctgg 600
ttcttctca tggctgtgat gtacatcact ggagctggcc ttatgtctgc tcgaattcct 660
gagcgttctt ttcttgaaa atttgacata tggttccagt ctcatcagat ttccatgtc 720
ctggtggtgg cagcagcctt tgtccacttc tatggagtct ccaaccttca ggaattccgt 780
tacggcctag aaggcggctg tactgatgac acccttctct gagccttccc acctgcgggg 840
tggaggagga acttcccaag tgccttttaa aataacttct ttgctgaagt gagaggaga 900
gtctgagttg tctgtttcta gaagaaacct cttagagaat tcagtaccaa ccaagcttca 960
gcccacttct acaccactg ggcaataaac ttccatttct catttctcta gctggggatg 1020
gggcatggct aaacttagcc atccccctct cagcaaggca tctaccggcc cctcacagag 1080
acagtaactt gaactcagc ttgagatttt acccttctct ccaaccattt tgggaaaatt 1140
atggactggg actcttcaga aattctgtct tttcttctgg aagaaaatgt ccttccctta 1200
cccccatct taactttgta tcttggttta taacaggcca tccatttttg tagcacactt 1260
ttcaaaaaca attatatacc ctggtcccat ctttctaggg cctggatctg cttatagagc 1320
aggaagaata aagccaccaa cttttaccta gcccggttaa tcatggaagt gtgtccaggc 1380

ttcaagtaac ttgagtttta attttttttt ttttcttggc agagtaatgt aaaattttaa	1440
tggggaaaaga tatttaatat ttaatactaa gctttaaaaa gaaacctgct atcattgcta	1500
tgtatcttga tgcaagact atgatgttaa taaaagaaag tacagaagac acttggcatt	1560
caaaaaaaaa aaaaaaaaaa aaa	1583

<210> 24
 <211> 1669
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (587)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1634)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1648)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1659)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1668)
 <223> n equals a,t,g, or c

<400> 24	
aggcgcttag gggctgaggg gcgatggcag gtgtcggggc tgggcctctg cgggcgatgg	60
ggcggcaggg cctgctgctt ctgcgctgt ggcgcacagg cgcccagggg ctctacttcc	120
acatcggcga gaccgagaag cgctgtttca tcgaggaaat ccccgacgag accatgggtca	180
tcggtcaggg gggctgaggg tggggaggcc ctttgtacct agctcagccc tcggcggcgc	240
tcctctctcc cgagcccagc cgggtcgctg gctccccag tacctagcct gagggtgccc	300
cgaggagccc agggcccctg cctagagctc cgggcccagc gtcggagggg gccgggcgga	360
gaggcggccc actaggccg gtcgtgacta tgtgtctgcc ccgcaggcaa ctatcgtacc	420
cagatgtggg ataagcagaa ggaggctctc ctgcccctga cccctggcct ggcatgcac	480
gtggaagtga aggaccccga cggcaagggt gtgctgtccc ggcagtagcg ctcgaggggc	540
cgcttcacgt tcacctccca cagcccgggt gaccatcaaa tctgtcngca ctccaattct	600
accaggatgg ctctcttcgc tgggtggcaaa ctgcgkgtgc atctcgacat ccagggtggg	660
gagcatgccca acaactaccc tgagattgct gcaaaagata agctgacgga gctacagctc	720
cgcgcccgcc agttgcttga tcaggtggaa cagattcaga aggagcagga ttaccaaagg	780
gcaagtgcac atctccttgt aatttgagag ggcagttgac ctttataccc actataccta	840
ctcaagtttc tgcttgggag atcagctctg cagagaatgg aatgagaagt attggtttag	900
ataggttgtt gtgttgggt ttttgagacg gaggtttact cttgttgccc agctggagt	960
gcaatgccat gatcttggct cactgcaacc tccgcctccc caggctgagg caggagaatg	1020
gcgtgagctc gggagggtgga gcttgcagtg agctgagatc gtgccactgc actccagcct	1080
gggcgacaga gtgagactcc ttctaaaaaa caaaaacaa accaaaacag tagttagggt	1140
acacacacac aaattctagt gattttcccc ccagtactac ccttgacttt tgaattctc	1200
gctttctcag agtttacaac atccttacca aacagccttc tccctcctta ccacaaaaaa	1260
araaaaaaa gttctggggg tgaggggaca ctccattctt aacatcctct attatcccag	1320
cccaattccc cagctctcac tgggactagt tgtacctatc ttcattcatt tgggtccagc	1380
atgactacct gttggtgcat gagctgatct ctccaaacct aacagccaga tgtagtctc	1440
tggtagctag atgctgggct gcattcagata ggatgcacag gatcatctcg ggaagcttgt	1500
tgacatagat tcctgtgcaa cacttcagat atagctctta ttagattttg tgttgggggt	1560
gtatggtagg tagaataatg ggcctaccac tgtgtaaca tatggataag tttacctaac	1620

atgcacagaag aganttaagt tgctaataatnag atgactgtgna aataaaatna

1669

<210> 25
<211> 1053
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1025)
<223> n equals a,t,g, or c

<400> 25
ctaggagcac cgagcagctt ggctaaaagt aaggggtgtcg tgctgatggc cctgtgacga 60
ctgacccgcg ctctgckctc tctgaacctg gcgccccga ccgtcgccgc ccctgccccg 120
agtctgttcc cgcgcgccca gatgatgaac aatggcctcc tccaacagcc ctctgccttg 180
atgttgctcc cctgcccggc agttcttact tctgtggccc ttaatgccaa ctttgtgtcc 240
tggaagagtc gtaccaagta caccattaca ccagtgaaga tgaggaagtc tgggggcccga 300
gaccacacag gccgaatccg ggtgcatggt attggcgggg gccacaagca acgttatcga 360
atgattgact ttctgcgttt ccggcctgag gagaccaagt caggaccctt tgaggagaag 420
gttatccaag tccgctatga tccctgtagg tcagcagaca tagctctggt tgctgggggc 480
agccggaaac gctggatcat cgccacagaa aacatgcagg ctggagatac aatcttgaac 540
tctaaccaca taggcgaat ggcatgtgct gctcgggaag gggatgcgca tctcttggg 600
gctctgcctg tggggaccct catcaacaac gtgaaagtgt agccaggccg gggtgcccaa 660
tatatccgag ctgcagggac gtgtggtgtg ctactgcgga aggtgaatgg cacagccatt 720
atccagctgc cctctaagag gcagatgcag gtgctggaaa cgtgcgtagc aacagttagc 780
cgagtatcca acgttgatca taacaaacgg gtcattggca aggcaggtcg caaccgctgg 840
ctgggcaaga ggcctaacag tgggcgggtg caccgcaagg ggggctgggc tggccgaaag 900
attcggccac taccctccat gaagagttac gtgaagctgc cttctgcttc tgcccaaaagc 960
tgatatccct gtactctaata aaaaatgccc cccccccgt taaaaaaaaa aaaaaaaaaa 1020
ctcgnggggg ggcccggtaa ccaattcggc cta 1053

<210> 26
<211> 1477
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (7)
<223> n equals a,t,g, or c

<400> 26
tgcaggnaac ggtccggaat tcccgggac aaacagtact gttgcacgtc gaattaagga 60
tctagctgct gacattgaag aagagcttgt ttgtagactg aaaatttgcg atgggttttc 120
actgcaacta gatgaatcag ctgatgtttc aggacttgct gtgctgcttg tgtttgttcg 180
ttataggttt aataagtcta ttgaggaaga cctactcctg tgtgaatctt tgcaaagtaa 240
tgctaccggt gaagaaatat tcaactgtat caacagtttt atgcagaaac atgaaattga 300
atgggaaaaa tgtgttgatg ttgttagtga tgcttctagg gcagtggatg ggaaaattgc 360
cgaagctgtc accttaataa aatatgtggc tcccgaagac accagtagtc actgcctatt 420
atacagacat gcaactggcag ttaaaataat gcctacatct ctaaaaaatg tgctagacca 480
ggcagtaaca atcatcaatt atattaaagc tcgaccacat caatccagac tattaataat 540
tttatgtgag gaaatgggtg ctcagcacac agcacttctt ctataacag aggtgagggtg 600
gctttctcga ggtaaagtgc ttgtaagact ttttgaactt cgtcgtgaac ttttggtttt 660
catggattct gcttttcgac tatctgattg ttttaacaaat tcatcttggc tgctaagact 720
tgcatatctt gcagatatatt ttactaaatt aaatgaagtt aatttgtcaa tgcaaggaaa 780
aaatgtgacc gtttttacag tatttgataa aatgtcgtca ttgttaagaa aattggaatt 840
ttgggcctca tctgtagaag aagaaaactt tgattgtttt cctacactca gtgatttttt 900
gactgaaatt aattctacag ttgataaaga tatttgcagt gccattgtgc agcacctaag 960
gggtttgcgc gctactctgt taaaataact tccgtgaaca aatgacaata atgcttgggt 1020
tagaaatcca ttacagtta ctgttaaac agcttcatta gtagcacggg actatgagag 1080
cctgattgat ttaacatctg attctcaagt gaagcaaaat ttttagtgaac tttcactaaa 1140

tgatttttgg	agtagcctaa	ttcaggaata	cccaagcatt	gcaaggcgtg	cagtgcgtgt	1200
actttcttct	tttgctacaa	tgacacctgt	tgaacacggg	ttttcatatt	acgctgcaac	1260
aaaaacaaaa	tataggaaaa	gacttgatgc	tgacacctcat	atgcgaatcc	gacttagcaa	1320
tattacacct	aatattaagc	ggatatgtga	taaaaagaca	caaaaacact	gttctcatta	1380
aaattggagg	agtttgcatt	tctcatgata	accaaagtga	agatgaaaa	aaaagatgat	1440
ttacttcaaa	aaaaaaaaaa	aaaaaaaaag	cgccgc			1477

<210> 27

<211> 2504

<212> DNA

<213> Homo sapiens

<400> 27

tcgaccacg	cgctccgag	tgctgcagg	actgggcctc	cttctccgc	ctggccatcc	60
ccagcatgct	catgctgtgc	atggagtgg	gggcctatga	ggtcgggagc	ttcctcagt	120
gcatcctcgg	catggtggag	ctgggcgctc	agtccatcgt	gtatgaactg	gccatcattg	180
tgtacatggt	ccctgcaggc	ttcagtgtgg	ctgccagtgt	ccgggttaga	aacgctctgg	240
gtgctggaga	catggagcag	gcacggaagt	cctctaccgt	ttccctgctg	attacagtgc	300
tctttgctgt	agccttcagt	gtcctgctgt	taagctgtaa	ggatcacgtg	gggtacattt	360
ttactaccga	ccgagacatc	attaatctgg	tggctcaggt	ggttccaatt	tatgctgttt	420
cccacctctt	tgaagctctt	gcttgacga	gtggtggtgt	tctgaggggg	agtggaaatc	480
agaaagtgg	agccattgtg	aataccattg	ggtamtatgt	ggttggcctc	cccatcgga	540
tcgcgctgat	gtttgcaacc	acacttggag	tgatgggtct	gtggtcaggg	atcatcatct	600
gtacagtctt	tcaagctgtg	tgttttctag	gctttattat	tcagctaaat	tggaaaaaag	660
cctgtcmgca	ggctcaggta	cacgccaatt	tgaagtaaaa	caacgtgcct	cggagtggga	720
attctgctct	ccctcaggat	ccgcttcacc	cagggtgccc	tgaaaacctt	gaaggatttt	780
taacgaacga	tggtgaaaag	acaggcgagc	ctcagtcaga	tcagcagatg	cgccaagaag	840
aacctttgcc	ggaacatcca	caggacggcg	ctaaattgtc	caggaacacg	ctggtgctgc	900
ggcgagggtc	tctgctctct	gggtctctct	taactctgct	gggtgggatt	ttagtgagat	960
tctatgtcag	aattcagtag	cgtggttaga	aagaaagtca	ggccaagtga	tgcttttgag	1020
cttacacaca	attcacaggc	ccaccagtga	caatttactg	tgagttaatg	tcattcaggt	1080
gtgcccattg	attttgaggg	ctggaaatgc	aaagacacat	ttttctataa	aaagaaaaag	1140
caactaaggt	taaaagctat	attgtggccc	aagacactgt	ctgaaagatg	acatgagtag	1200
taattcacca	caactctgaac	caagcaagga	tcaattgtgt	gactgcattg	gccaatgggt	1260
ttgatacttc	tgctatTTTT	ttagacacaa	acccataaac	taactgctta	agaattcata	1320
ctgcttgaat	tatgtaaaat	atattttaca	gtatatcttt	ccttgggcct	tagattacta	1380
ttcactgggc	aaatgggtatt	tgTTTTTgtt	ttatTTTTTT	TTTTaataga	cggaagtctt	1440
gctctgtcat	gcaggctgga	gtgcggtggt	gcgatcatag	ctcactgcag	cctcgaactc	1500
ttgggcttca	agcaatcttc	ctgtgtcagc	caccagagta	gctgagacta	caggggtatg	1560
ccaccatgcc	cagctggcat	ttgttaatct	tcatttgagg	tctagatcta	ggcactgtgg	1620
acactgaaaa	acagtgggga	aatcttttca	gctgtggaaa	tccaaacaaa	gactgataat	1680
tctgtgtarg	ggtgtgtgcs	tgacgtactg	carcctyaam	ctyctgggct	yaagtgatcc	1740
tcccacctca	gcctcctgag	tagctgagac	cacaggcgtg	tgccaccacg	cctagctaat	1800
ttttwawacc	rgggtcwamc	ctttgtttcc	caggstggtg	ttgaattcct	gggatcaagc	1860
aatycttcca	cctkgsmtc	ccaaagtgtt	gggattatag	gcatgagcca	ccasgactgg	1920
ccagaggaca	aaattttaat	aaaggctcta	gcttaagcag	taatcytact	tcattaagcc	1980
ttcctgggg	gcggtacaca	ccgttaattc	agcaaccctc	agtaataact	aagtatgctc	2040
agtgtgtgta	aagtggatta	caccaaatta	agtcattctt	atcacacca	atcaaaagtc	2100
aagaagccag	ggataaaagc	acctcaggca	cataacatta	atctagtaat	gtaattctct	2160
gcacatccag	ctgggtgaac	tgctgtgctg	aagctgggac	cagctttgtc	cataactgct	2220
gagagaactt	gctgaagctc	taggaataat	tttgctgcc	cggttgctca	ccagttgtag	2280
cttgccagct	cccaacaccc	ttcctgtgtc	caataaactt	tctcaaagag	caatuu .gac	2340
atttcttttg	ataaaacctc	cagccttctc	tgtgtgttgc	cgacataacc	aggaccaact	2400
ggtctacatg	gatgccctga	acatgcaatt	cttcttcca	aaataaaaca	ttaatagag	2460
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaagggc	ggcc		2504

<210> 28

<211> 1866

<212> DNA

<213> Homo sapiens

<400> 28


```

ggcacgagaa tacatacgat ccttgtctac caggagtcta atagaaagat ggacagcggtg 60
gacctgcca gcagccaggc catggagctc tctgatgtca cctcattga ggggtgtgggt 120
aatgaggtga tgggtgtggc aggtgtggtg gtgctgattc tagccttggg cctagcttgg 180
ctctctacct acgtagcaga cagcggtagc aaccagctcc tgggcgctat tgtgtcagca 240
ggcgacacat ccgtcctcca cctggggcat gtggaccacc tgggtggcagg ccaaggcaac 300
cccagacca ctgaactccc ccatccatca gagggtaatg atgagaaggc tgaagaggcg 360
gggtgaaggta ggggagactc cactggggag gctggagctg ggggtgtgtg tgagccagc 420
cttgagcatc tccttgacat ccaaggcctg cccaaaagac aagcaggtgc aggcagcagc 480
agtccagagg ccccccctgag atctgaggat agcacctgcc tccctcccag ccctggcctc 540
atcactgtgc ggctcaaatt cctcatgat accgaggagc tggctgtggc taggcccagag 600
gataccgtgg gtgcctgaa gagcaaatc tccctggac aagaaagcca gatgaaactg 660
atctaccagg gccgcctgct acaagacca gccgcacac tgcgttctct gaacattacc 720
gacaactgtg tgattcactg ccaccgctca ccccagggt cagctgttcc aggccctca 780
gcctccttgg cccctcggc cactgagcca ccagccttg gtgtcaatgt gggcagcctc 840
atgggtcctg tctttgtggt gctgttgggt gtggtctggt acttccgaat caattaccgc 900
caattcttca cagcacctgc cactgtctcc ctgggtggag tcaccgtctt cttcagcttc 960
ctagtatttg gtagtatgg acgataagga cataggaaga aaatgaaagg catggtcttt 1020
ctcctttatg gcctcccac tttcctggc cagagctggg cccaagggcc ggggaggagg 1080
gggtggaaag gatgtgatg aaatctctc cataggacac aggaggcaag tatgcggcct 1140
ccccttctca tccacaggag tacagatgtc cctcccgtgc gagcacaact caggtagaaa 1200
tgaggatgtc atcttcttc acttttaggg tcctctgaag gaggtaaaag ctgctggcca 1260
agctcagtg ggagcctggg ctctgagatt cctcccacc tgggttctg actcttcca 1320
gtgtcctgca tgtctgccc cagcaccag ggtgcctgc aagggcagct cagcatggcc 1380
ccagcacaac tccgtaggga gcctggagta tccttccatt tctcagccaa atactcatct 1440
tttgagactg aaatcacact ggcggaatg aagattgtgc cagccttctc ttatgggac 1500
ctagccgctc tcacctctt cctctacccc ttagcaggaa tagggtgtcc tccctcttt 1560
caaagcactt tgcttgcat ttattttatt tttttaagag tccttcatag agctcagta 1620
ggaaggggat ggggcaccaa gccaagccc cagcattggg agcgccagg ccacagctgc 1680
tgctcccgtg gtcctcaggc tgtaagcaag agacagcact ggcccttggc cagcgtcta 1740
ccctgcccac ctccaaggac tgggtatgga ttgctgggct ctaggctctt gcttctggg 1800
ctattggagg gtcagtgtct gtgactgaat aaagtccat tttgtggtaa aaaaaaaaaa 1860
aaaaaa 1866

```

<210> 29
 <211> 1501
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (434)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (441)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1300)
 <223> n equals a,t,g, or c

```

<400> 29
ggacagcggt atcagcctgc tgggtgaataa cgccggtgtg ggccgccagg cttegtgtgt 60
ggagtcggat gccgacaaaa tggacgcgat gattctgtct aacgtactgg cgtgacccg 120
cctggccaaa gccgcggcaa ccaactttgt cgcccagggc cgtggcacga tcatcaacat 180
cggctcgatt gtcgtctctg ctcccaaagt gctgaacggc gtgtatggcg gtaccaaaag 240
gttcgtgcag gcgttcagcg aatcgtgtga gcatgagctg agtgacaagg gcgtagtgg 300
ccaggtgtgt ctgccaggcg ctaccgccac ggagttcttg gacatcggcg gcctgcctgt 360
gaaacaacct gccggaagcc atggtgatga ccaccgaaa cctggtggac gccgcctg 420
caggccttgc ccanggcraa ncgtgacgat tccgtccctg ccggacagcg cagattggga 480
cactacgaac gcgcgaggct ggccctgggt ccgaacctgt cgcaccgtga acccgccgct 540

```

```

cgttatgggt tgaagtaatc cggactagcg cagccggggt taaacgcagg cttcctgatt 600
gcctgggagg cctgttcata cccgtaggcg accgacagca acgtggcttc gctcaaattt 660
ttcccataga agtgaacggc tgctggcatc ccttcgtcgt ccatgcccga tggatggag 720
ataccgggat aaccggccac cgccgagtag tagtaactgt atgagtgaat gttggacatc 780
attgcatcaa gcttatgctc ggccagcggc ttatcgatgg tgcttttgaa aatcggggccg 840
atggcagccc ataactcatt gcgcgcctca tcatcgatat ccatcccgtt gatcatggtg 900
agcatctggt gatccggcac acccgggaccg ctgttgcgct cgttgaattc aatcagctca 960
gccagcgact tcaccggcaa gcctgcccgt ccggccagggt aggcctcaag ctgggtgttta 1020
acgtccgata acaacgcgtc gttatatgtt tcatgggttt cgtacgggac gccctcacc 1080
agttgaccca cgggtaccaa tgcgcgccc ttgctcgca gcaacgtaat ggcacccctc 1140
aagtgcstct gtcggctttt ttccgggggt cgttgccatc ttctacagat aactcaggca 1200
acggcgatata accgatgcgc ttgcccacca agcgctcagg cttgattccc tgggtgttagc 1260
ggttgggtatc gtcacgcgca tccagtgcct gcgcgcgatn acgcacgtta cgggtgaagg 1320
tgcccaccgt gtccctggcg gaactgggtc tacccttcg gtactcacta atccttcggt 1380
cggtttgaaa ccaataaacac cggttgtaagc cgccggcgta atgattgaac cattgggttc 1440
gacccccaat gccaaaggga caatcccttg tgcaacggct accgcagagc ccgtactcga 1500
g 1501

```

```

<210> 30
<211> 1752
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1099)
<223> n equals a,t,g, or c

```

```

<400> 30
aaggtagccc tgcaggtagc ggtccggaat tcccggtcgc acccagcgt ccgtccagga 60
cagagagtgc acaaaactacc cagcacagcc ccctccgccc cctctggagg ctgaagaggg 120
attccagccc ctgccaccca cagacacggg ctgactgggg tgcctgcccc ccttgggggg 180
gggcagcaca gggcctcagg cctgggtgcc acctggcacc tagaagatgc ctgtgccctg 240
gttcttgcgt tccttggcac tgggcccgaag cccagtggtc ctttctctgg agaggcttgt 300
ggggcctcag gacgctaccc actgctctcc gggcctctcc tgccgcctct gggacagtga 360
catactctgc ctgcccgggg acatcgtgcc tgcctcgggc ccctgtctgg cgctacgca 420
cctgcagaca gagctggtgc tgaggtgcca gaaggagacc gactgtgacc tctgtctgcg 480
tgtggmtgtc cacttggccg tgcattggga ctgggaagag cctgaagatg aggaaaagt 540
tggaggagca gctgacttag ggggtggagg gccatggaat gcctctctcc aggcccaagt 600
cgtgctctcc ttccaggcct accctactgc ccgctgcgtc ctgctggagg tgcaagtgcc 660
tgctgcctct gtgcagtttg gtcagtctgt gggctctgtg gtatatgact gcttcgaggc 720
tgccctaggg agtgaggtag gaattctggt ctatactcag ccaggtacg agaaggaaat 780
caaccacaca cagcagctgc ctgactgcag ggggctcgaa gtctggaaca gcatcccgag 840
ctgctggggc ctgccctggc tcaacgtgct agcagatggt gacaacgtgc atctggttct 900
gaatgtctct gaggagcagc acttcggcct ctccctgtac tggaaatcagg tccaggggcc 960
cccaaaaccc cggtaggcaca aaaacctgac tggaccgcag atcattacct tgaaccacac 1020
agacctggtt cctgcctctt gtattcagggt gtggcctctg gaacctgact ccgttagacg 1080
aacatctgcc ccttcaggna ggaccccgcc gcacaccaga acctctggca agccgcccga 1140
ctgagactgc tgaccctgca gagctggctg ctggacgcac cgtgctcgtt gcccgcagaa 1200
gcggcactgt gctggcgggc tccgggtggg gacccctgcc agccactggt cccacgcgtt 1260
tcctgggaga aygtcactgt ggacaagggt ctgcagttcc cattgctgaa aggccaccct 1320
aacctctgtg ttcaggtaga cagctcggag aagctgcagc tgcaggagtg cttgtggg 1380
gactccctgg ggcctctcaa agacgatgtg ctactgttgg agacacgagg ccccaggac 1440
aacagatccc tctgtgcctt ggaacccagt ggctgtactt cactaccag caaagcctcc 1500
acgagggcag ctgccttgg agagtactta ctacaagacc tgcagtcagg ccagtgtctg 1560
cagctatggg acgatgactt gggagcgcta tgggcctgcc ccatggacaa atacatccac 1620
aagcgctggg ccctcgtgtg gctggcctgc ctactctttg cctgcgcttt ccctcatcct 1680
ccttctcaaa aaggatcacg cgaaagggtg gctgaggctc ttgaaacagg acgtccgctc 1740
ggggggcgcc gc 1752

```

```

<210> 31
<211> 2152

```

<212> DNA

<213> Homo sapiens

<400> 31

ccgctttgtt	ctccagatgt	gaatagctcc	actataccag	cctcgtcttc	cttccggggg	60
acaaacgtggg	tcaggggcaca	gagagatatt	taatgtcacc	ctcttggggc	tttcatggga	120
ctccctctgc	cacatttttt	ggaggttggg	aaagtgtcta	gaggcttcag	aactccagcc	180
taatggatcc	caaactcggg	agaatggctg	cgctccctgt	ggctgtgctg	ctgctgctgc	240
tgctggagcg	cggcatgttc	tcctcaccct	ccccgcccc	ggcgtgttta	gagaaagtct	300
tcagatcat	tgacctccat	caggatgaat	ttgtgcagac	gctgaaggag	tgggtggcca	360
tcgagagcga	ctctgtccag	cctgtgcctc	gcttcagaca	agagctcttc	agaatgatgg	420
ccgtggctgc	ggacacgctg	cagcgcctgg	gggcccgtgt	ggcctcggtg	gacatgggtc	480
ctcagcagct	gcccgatggg	cagagtcttc	caatacctcc	cgctcatctg	gccgaactgg	540
ggagcgatcc	cacgaaaggc	accgtgtgct	tctacggcca	cttggacgtg	cagcctgctg	600
accggggcga	tgggtggctc	acggaccctt	atgtgctgac	ggaggtagac	gggaaacttt	660
atggacgagg	agcgaccgac	aacaaaggcc	ctgtcttggc	ttggatcaat	gctgtgagcg	720
ccttcagagc	cctggagcaa	gatcttcctg	tgaatatcaa	attcatcatt	gaggggatgg	780
aagaggtctg	ctctgttgcc	ctggaggaaac	ttgtggaaaa	agaaaaggac	cgattcttct	840
ctggtgtgga	ctacattgta	atttcagata	acctgtggat	cagccaaagg	aagccagcaa	900
tcacttatgg	aacccggggg	aacagctact	tcattggtgga	ggtgaaatgc	agagaccagg	960
attttcactc	aggaaccttt	ggtggcatcc	ttcatgaacc	aatggctgat	ctggtgtgct	1020
ttctcggtag	cctggtagac	tcgtctgggt	atatcctggt	ccctggaatc	tatgatgaag	1080
tggttctctt	tacagaagag	gaaataaata	catacaaaagc	catccatcta	gacctagaag	1140
aataccggaa	tagcagccgg	gttgagaaat	ttctgttcga	tactaaggag	gagattctaa	1200
tgcacctctg	gaggtaccca	tctctttcta	ttcatgggat	cgagggcgcg	tttgatgagc	1260
ctggaaactaa	aacagtcata	cctggccgag	ttataggaaa	atcttcaatc	cgtctagtcc	1320
ctcacatgaa	tgtgtctgcg	gtggaaaaac	aggtgacacg	acatcttgaa	gatgtgttct	1380
ccaaaagaaa	tagttccaac	aagatggttg	tttccatgac	tctaggacta	caccctggga	1440
ttgcaaatat	tgatgacacc	cagtatctcg	cagcaaaaag	agcgatcaga	acagtgtttg	1500
gaacagaacc	agatatgac	cgggatggat	ccaccattcc	aattgccaaa	atgttccagg	1560
agatcgctca	caagagcgtg	gtgctaattc	cgctggggagc	tgttgatgat	ggagaacatt	1620
cgcagaatga	gaaaatcaac	aggtggaact	acatagaggg	aaccaaatat	tttgcctgct	1680
ttttcttaga	gatggcccag	ctccattaat	cacaagaacc	ttctagtctg	atctgatcca	1740
ctgacagatt	cacctcccc	acatccctag	acagggatgg	aatgtaaata	tccagagaat	1800
ttgggtctag	tatagtacat	tttcccttcc	atttaaaatg	tcttgggata	tctggatcag	1860
taataaaata	tttcaaaggc	acagatgttg	gaaatgggtt	aaggtccccc	actgcacacc	1920
ttcctcaagt	catagctgct	tgacagcaat	tgatttcccc	aagtcctgtg	caatagcccc	1980
aggattggat	tccttccaac	cttttagcat	atctccaacc	ttgcaatttg	attggcataa	2040
tcactccggt	ttgctttcta	ggtcctcaag	tgctcgtgac	acataatcat	tccatccaat	2100
gatcgccctt	gctttaccay	tctttccttt	tatcttatta	ataaaaatgt	tg	2152

<210> 32

<211> 1757

<212> DNA

<213> Homo sapiens

<400> 32

aggctttcca	cccagaccgt	caacttcggg	acagtggggg	agacgggtcac	ccttcacatc	60
tgcccagaca	gggatgggga	tgaggcggca	cagcctgatg	ctgctgccat	ggtggcttgg	120
ggcagcgggg	agaaaggagt	gtcacaggga	gcagctcgtg	gctgcagtgg	aagtcactga	180
gcaagagact	aaagtcccca	agaaaaccgt	catcatagaa	gagaccatca	ccactgtggt	240
gaagagccca	cgtggccaac	gacggtycce	cagcaagtcc	ccctcccgct	caccttccc	300
ctgctctgcc	agcccgtgga	ggccaggcct	actggcccc	gacctgctgt	acctgccagg	360
tgctggccag	ccccgcaggc	cggargcaga	accaggccag	aagccertgg	tgcccacact	420
gtatgtgacg	gaggccgagg	cccactctcc	agctctgccc	ggactctcgg	ggccccagcc	480
caagtgggtg	gaggtggagg	agaccattga	agtcgggtg	aagaagatgg	gcccgcaggg	540
tgtgtctccc	accacagagg	tgcccaggag	ctcatcgggg	catctcttca	cactgcccgg	600
tgcgaccccc	ggagggggacc	ccaattccaa	caactccaac	aacaagctgc	tggcccagga	660
ggcctggggc	cagggcacag	ccatggtcgg	cgctcagagag	ccccttgtct	tccgcgtgga	720
tgccagaggc	agtgtggact	gggctgcttc	tggcatgggc	agcctggagg	aggagggcac	780
catggaggag	gcgggagagg	aagaggggga	agacggagac	gcctttgtga	cggaggagtc	840
ccaggacaca	cacagccttg	gggatcgtga	ccccaaagatc	ctcacgcaca	acggccgcat	900
gctgacactg	gctgacctgg	aagattacgt	gcctggggaa	ggggagacct	tccactgtgg	960

tgccctggg	cctggcgccc	ctgatgaccc	tccctgagag	gtctcggtga	tccagagaga	1020
gatcggggag	cccacgggtg	gcagcctgtg	ctgctcagcg	tggggcatgc	actgggtccc	1080
cgaggccctc	tgggcctctt	taggcctgag	ccccgtggg	cgtcaccacc	gggaccccag	1140
gtccgtagcc	ttgagggcac	ctccttcctc	ttgcgggagg	ccccggctcg	gcctgtgggc	1200
agtgtccct	ggacgcagtc	tttctgcacc	cgcatccggc	gttctgcgga	cagtggccag	1260
agcagcttca	ccacagagct	ttccacccag	accgtcaact	tggggacagt	gggggagacg	1320
gtcacccctc	acatctgtcs	ctggccwccg	gccttcttac	ctcactcaac	ttcagccagg	1380
aggactgggt	gggtgcttga	atgttggaat	gaccggctca	aagacctcag	ctctgggctg	1440
tttctgtca	gcctggcagg	agcctcagga	ctgtggacga	aggatgtggc	cttgggcatt	1500
tgtcctgttc	ccacatgggc	ctggtccctc	cctcctggcc	ccagccacag	ctgccaggcc	1560
tgacatggcc	ttgcctctcc	tgcagtcttg	gtgactgaga	cccttgggtg	gcgcttccca	1620
gctctgggc	ccctcctggc	ctttcttgca	gggtggacac	agggtctgtg	tgtgggcagc	1680
agccctgtc	tctcagcaag	aataaagcag	cttctgtg	aaaaaaaaa	aaaaaaaaa	1740
aactcgagcg	gcacgag					1757

<210> 33
 <211> 1466
 <212> DNA
 <213> Homo sapiens

<400> 33						
ggcacaggct	gggacttttg	gctggctgca	gtctgtctga	gggcggccga	agtggctggc	60
tcatttaaga	tgaggcttct	gctgcttctc	ctagtggcgg	cgtctgcgat	ggtccggagc	120
gaggcctcgg	ccaatctggg	cggcgtgccc	agcaagagat	taaagatgca	gtacgccacg	180
gggcccgtgc	tcaagttcca	gatttgtgtt	tcctgagggt	ataggcgggt	gtttgaggag	240
tacatgccc	ttattagcca	gcggtaccca	gacatccgca	ttgaaggaga	gaattacctc	300
cctcaaccaa	tatatagaca	catagcatct	ttcctgtcag	tcttcaaact	agtattaata	360
ggcttaataa	ttgttgga	ggatcctttt	gctttctttg	gcattgcaagc	tcctagcatc	420
tggcagtggg	gccaagaaaa	taagggttat	gcattgtatga	tgggttttctt	cttgagcaac	480
atgattgaga	accagtgtat	gtcaacaggt	gcatttgaga	taactttaaa	tgatgtacct	540
gtgtgggtca	agctggaatc	tggtcacctt	ccatccatgc	aacaacttgt	tcaaattctt	600
gacaatgaaa	tgaagctcaa	tgtgcataatg	gattcaatcc	cacaccatcg	atcatagcac	660
cacctatcag	cactgaaaac	tcttttgcat	taagggatca	ttgcaagagc	agcgtgactg	720
acattatgaa	ggcctgtact	gaagacagca	agctgttagt	acagaccaga	tgctttcttg	780
gcaggctcgt	tgtacctctt	ggaacacctc	aatgcaagat	agtgtttcag	tgctggcata	840
ttttggaatt	ctgcacattc	atggagtgca	ataatactgt	atagctttcc	ccacctccca	900
caaaatcacc	cagttaaatgt	gtgtgtgtgt	ttttttttta	aggtaaacat	tactacttgt	960
aacttttttt	cttagtcata	tttgaaaaag	tagaaaaattg	agttacaatt	tgattttttt	1020
tccaaagatg	tctgttaaat	ctgttggtgt	tttatatgaa	tattgttttt	ttatagttaa	1080
aaattgatcc	tttggaatc	cagttgaagt	tcccaaatat	tttataagag	tttatcagac	1140
atctctaatt	tgcccatgtc	cagttttatac	agtttacaaa	atatagcaga	tgcaagatta	1200
tgggggaaat	cctatatcca	gagtactcta	taaatttttg	tgtatgtgtg	tatgtgcgtg	1260
tgattaccag	agaactacta	aaaaaaccaa	ctgcttttta	aatcctattg	tgtagttaaa	1320
gtgtcatgcc	ttgaccaatc	taatgaattg	attaattaac	tgggccttta	tacttaacta	1380
aaaaaaaaac	taagcagata	tgagttaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa	actcga				1466

<210> 34
 <211> 526
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (283)
 <223> n equals a,t,g, or c

<400> 34						
ggggacgtgc	acggggccgc	cctcctggcc	ctgaagctgc	gccggcctcc	ctgagcgttt	60
cgctgcggag	ggaagtccac	tctcggggag	agatgctgat	gccggtcac	ttcctgctgc	120
tcctgctgct	gctcctgggg	ggccccagga	caggcctccc	ccacaagttc	tacaaagcca	180
agcccatctt	cagctgcctc	aacaccgccc	tgtctgaggc	tgagaagggc	cagtgggagg	240

atgcatccct	gctgagcaag	aggagcttcc	actacctgcy	canaagsacg	cctcttcggg	300
agaggaggag	gagggcaaa	agaaaaagac	tttccccatc	tctggggcca	ggggtggarc	360
cagaggcacc	cggtacagat	acgtgtccca	agcacagccc	aggggaaagc	cacgccagga	420
cacggccaag	agtccccacc	gcaccaagtt	caccctgtcc	ctcgacgtcc	ccaccaacat	480
catgaacctc	ctcttcaaca	tcgccaaggc	caagaactgc	gtgccc		526

<210> 35
 <211> 2412
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (329)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (340)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (977)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1117)
 <223> n equals a,t,g, or c

<400> 35						
cacgagtttt	aaatcaattt	tttttcaagc	aatcagattc	ttttctccta	gaggagctgt	60
gggcaagaaa	actaatgaat	tctacatcct	tctcatcacc	tggttttaaa	tgttttctgc	120
tctgagtaaa	cagtaattac	tgtttaagta	catctcagca	gaattttatc	ccaattgcaa	180
cagttcatgt	tctctcta	gtaatctctg	cggaggaaat	gatcgtcaag	ggaagcaggc	240
tgacctgtct	acgggatggc	gttcttaca	tctgcatctt	atgtaatggt	gattctgtgt	300
gcctgtgtca	taattattgg	aatattatnt	tatgcttttn	tttttgagac	tctatctcca	360
aaaaaaagaa	gagacataga	aatttgaaga	aggatccctt	aatgggtctac	accgtcttcc	420
aaagtcaaga	agtggcagct	gatattccatt	tgaaagtaga	atcctagctt	ttcagagcta	480
gacmaggcct	cagaaactat	agttgaattc	ctcattgtac	caatgagaaa	ctcaggccta	540
gatgggttaa	aagaggtgtg	ttgtagcagt	gctgggacag	atctcggttt	ttctgcttcc	600
tatacaatcc	tcttcaaccc	aatactacaa	tgtatttatt	atcacatatt	aagctggaga	660
ttttagacca	tggtattaga	gttgcaactg	tttatcctat	agattccagc	cacattttaa	720
acacataact	tcatgtagtt	aggccactaa	aaataaagta	atccatcaaa	ctagtaatac	780
actagagaat	ttgacctaca	tactaagatg	cctgaaatcc	acagtatatg	gcaattttaac	840
ccccatctaa	tagtggctca	atcaagtagc	taaacatatt	tatttcaact	agatgggttg	900
ttgtttggta	gaaggaatgg	actccttggg	ctattttggg	aacaaaaaag	gactaggaca	960
caaatcaaag	ccacatncac	agtaagaaat	ccgggctgat	ctctgccaa	aaaagktaca	1020
aagaataatt	acttgatcac	gtggggaaat	ttcgacataa	aagaagtaat	ggataaaaarg	1080
aaagaaaaat	gaccaattgc	tgragmcaat	aattatngca	accctaaacc	agaaagcact	1140
aaagccaggaa	gtcaaaaact	aagtcatmca	catatgacaa	ggtgcggggg	ttggtcctga	1200
gacttcagtg	agaatatgtc	cgatcaggat	atgcaaaagaa	ccatttggaa	gatttctagt	1260
tcataaggga	agtaccaaat	gaagtggatg	ggaccatacg	caatttgcac	aggacccccca	1320
aggaggaaat	agtatgacat	ggtagtaaaa	aagcaatcac	gactacactc	acaatttttag	1380
gagaaaaataa	aactaaatcc	agaatttgaa	gccaacaaca	acaaaaaaag	tcattatttta	1440
gggtatacgt	tctgtgggc	agtaccttgc	aaagtagaac	atcttcaaga	agaaatatttt	1500
gacttgaggt	aaggctttca	agattgccat	attacattca	taaaggcaaa	ctcatccttg	1560
agaccaaaat	gcagaagat	tagaaattaa	ggctttgttt	taaagaaatg	ttgacatcat	1620
actggaaatt	attatccagc	tacttacaca	ttcgttttta	aatccatccc	tatgttttagc	1680
tgccaaaatg	caaactgcgc	attgywctct	cacggagagc	gccacaggtc	accagctatt	1740
attctcccag	gagtcattga	gtaggctgcc	ccaagtacca	cataggaaac	tcaacgaact	1800
attttcattt	caaggaccat	tagaaacaga	aaggaaaaa	gaaggtcagg	gaaacttagt	1860

ttctaacaaa	ggaagtgagg	cactttgaaa	aagaaaatat	ttagagaacg	gagaggaagc	1920
taaacccaaa	caaccaaaac	gcacagctga	caattattcc	gggaagtgg	taacttctgc	1980
ctggtctcta	gaagcacagg	aagaaaggac	tgtagcgtg	aagaacactc	cagggttctg	2040
ggtatctagg	cagagtcagt	caacagggct	aaccatgtga	taatcctggg	taattccacc	2100
tcacagttca	ctaaaaaaca	agcggaaacc	tgggcaaagc	cctttggggc	tttaaatagca	2160
atggaggaca	tcaccctgtc	actttctctg	cttctacaca	gcaggcaatc	aaggaaaact	2220
tgccaagaaa	tatgagtga	taaatgattt	tgaaagtttc	attgagcagg	aacatgaaaa	2280
ggatgatttg	gggatagctg	gaaggatagt	tacttgcatt	aataatattt	attcaccgtc	2340
agtgatgat	ttctcaatag	aaagattgta	tttaaaatgt	acaactacaa	aaaaaaaaaa	2400
aaaaaactcg	ag					2412

<210> 36
 <211> 1274
 <212> DNA
 <213> Homo sapiens

<400> 36	
ggcacgagga	aagaccaact
gggagaggat	tacacagtac
ggttagtttc	actatctgct
cttctgaatt	actttgcctg
ttattaccct	tggtcaccct
gctcatgcag	ctgatagaat
tgctcttttc	tcaacatggc
cagccctacc	acccccacc
tgctcagttct	gccattcacg
ttctcatatg	cctaattggaa
tcccactgac	tcaatagaac
acttgcatgg	gccaatgaga
tgactgggtt	agcttggcct
gccactgctc	attcagcctg
atccacagcc	cagagcagag
gtctattggt	ataaaccagt
aaatttggag	aatgacacat
gggttggggc	cagcaatctg
ggcttggaat	ccactgttat
gctgactaat	acatctaccc
aagctcctgg	aatgaggttc
atctgatcct	cgag

<210> 37
 <211> 1036
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (43)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (47)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (58)
 <223> n equals a,t,g, or c

<400> 37	
caaaccat	ttacgtgcac
actgatacac	cgtcacgct
gcnggtnacc	ggtccggnat
	60

tccccgggtcg	acccacgcgt	ccgggggaag	caagcactta	tttggctact	tgggtgtccat	120
ggggaaagaa	ttcctaatac	tccttatgtg	ttagaggact	ttgttgagaa	tgtgaagtcg	180
gaaacatttc	cagctgttaa	gatggagctg	ctcactgctt	tgtgcgcct	ttcctctcc	240
cgacctgctg	agtgccagga	catgctagga	cgtttgttgt	attactgcat	agaggaagaa	300
aaagatatgg	ctgtacggga	ccgagggtct	ttctattatc	gcctcctctt	agttggcatt	360
gatgaagtta	agcggattct	gtgtagccct	aaatctgacc	ctactcttgg	acttttggag	420
gatccggcag	aaagacctgt	gaatagctgg	gcctcagact	tcaacacact	ggtgccagtg	480
tatggcaaa	cccactgggc	aactatctct	aaatgccagg	gggcagagcg	ttgtgaccca	540
gagcttctta	aaacttcatc	ctttgccgca	tcaggaccct	tgattcctga	agagaacaag	600
gagagggtac	aagaactccc	tgattctgga	gccctcatgc	tagtcccca	tcgccagctt	660
actgctgatt	attttgagaa	aacttggctt	agccttaaag	ttgtctcatca	gcaagtgttg	720
ccttggcggg	gagaattcca	tcctgacacc	ctccagatgg	ctcttcaagt	agtgaacatc	780
cagaccatcg	caatgagtag	ggctgggtct	cggccatgga	aagcatacct	cagtgtctag	840
gatgatactg	gctgtctgtt	cttaacagaa	ctgctattgg	agcctggaaa	ctcagaaatg	900
cagatctctg	tgaaacaaaa	tgaagcaaga	acggagacgc	tgaatagttt	tatttctgta	960
ttagaaactg	tgattggaac	aattgaagaa	ataaaatcat	aacagagaaa	aaaaaaaaaa	1020
aaaaaagggc	ggccgc					1036

<210> 38
 <211> 1379
 <212> DNA
 <213> Homo sapiens

<400> 38						
gcggcgcggg	tgggggttgt	gcgttttacg	caggctgtgg	cagcgacgcg	gtccccagcc	60
tgggtaagaa	tggccccatg	gccccgaag	gcctagtcct	agctgtgtct	tggggcctca	120
gcctcttctt	caacctccca	ggacctatct	ggctccagcc	ctctccacct	ccccagctct	180
ctcccccgcc	tcagcccat	ccgtgtcata	cctgccgggg	actggttgac	agctttaaca	240
agggcctgga	gagaaccatc	cgggacaact	ttggaggtgg	aaacactgcc	tgggaggaag	300
agaatttgtc	caaatacaaa	gacagtgaga	ccgcctgggt	agaggtgctg	gaggggtgtg	360
gcagcaagtc	agacttcgag	tgccaccgcc	tgctggagct	gagtgaggag	ctggtggaga	420
gctggtgtgt	tcacaagcag	caggaggccc	cggacctctt	ccagtggctg	tgctcagatt	480
ccctgaagct	ctgctgcccc	gcaggcacct	tcgggcccct	ctgccttccc	tgctctgggg	540
gaacagagag	gccctgcggt	ggctacgggc	agtgtgaagg	agaagggaca	cgagggggca	600
gcggggcactg	tgactgccaa	gccggctacg	ggggtgaggc	ctgtggccag	tgtggccttg	660
gtctacttga	ggcagaacgc	aacgccagcc	atctgggtatg	ttcggcttgt	tttggccccct	720
gtgccccgatg	ctcaggacct	gaggaatcaa	actgtttgca	atgcaagaag	ggctggggccc	780
tgcatcacct	caagtgtgta	gacattgatg	agtgtggcac	agagggagcc	aactgtggag	840
ctgaccaat	ctgcgtgaac	actgagggct	cctatgagtg	ccgagactgt	gccaaggcct	900
gcttagctg	catgggggca	gggccaggtc	gctgtaagaa	gtgtagccct	ggctatcagc	960
aggtgggctc	caagtgtctc	gatgtggatg	agtgtgagac	agaggtgtgt	ccgggagaga	1020
acaagcagtg	tgaaaacacc	gaggggcggt	atcgctgcat	ctgtgccgag	ggctacaagc	1080
agatggaagg	catctgtgtg	aaggagcaga	tcccaggtgc	attccccatc	tttaactgatt	1140
taacccttga	aacaaccga	cgctggaagt	tgggttctca	tccccactct	acatatgtaa	1200
aaatgaagat	gcagagagat	gaagctactt	tcccagggtc	atatggcaag	caagtgcgaa	1260
agctggggtc	ccaatccaga	cagcttgacc	gtggaacgag	actcatacac	gtaataaatg	1320
ctctgcccc	aacttgtcca	ccacaaaaaa	aaaaaaaaaa	aaaaaaaaaa	ggcggccgc	1379

<210> 39
 <211> 1932
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (293)
 <223> n equals a,t,g, or c

<400> 39						
ggcacgaggg	cgccctgggt	gtcagcgggt	cggctcccgc	gcacgtcccg	gccgtcgcgc	60
asctcggcac	ctgcaggctc	gtgcgtcccg	cggctggcgc	ccctgaactc	gtcccggcca	120
gggaggggcca	tgatttccct	cccggggccc	ctgggtgacca	acttgcgtgcg	gtttttgttc	180

ctggggctga	gtgccctcgc	gccccctcgc	cgggccagc	tgcaactgca	cttgcccgc	240
aaccgggtgc	agggcggtgga	gggaggggaa	gtggtgcttc	cagcggtgga	cancttgac	300
ggggagggtgt	cttcacacca	gcatggggag	gtgccctttg	tgatgtggtt	cttcaaacag	360
aaagaaaagg	aggatcaggt	gttgctctac	atcaatgggg	tcacaacaag	caaacctgga	420
gtatccttgg	tctactccat	gcccctcccg	aacctgtccc	tgcggtgga	gggtctccag	480
gagaaagact	ctggccccta	cagctgtctc	gtgaatgtgc	aagacaaaca	aggcaaatct	540
agggggccaca	gcacaaaac	cttagaactc	aatgtactgg	ttcctccagc	tcctccatcc	600
tgccgtcttc	aggggtgtgcc	ccatgtgggg	gcaaacgtga	ccctgagctg	ccagtctcca	660
aggagtaagc	ccgctgtcca	ataccagtgg	gatcggcagc	ttccatcctt	ccagactttc	720
tttgaccag	cattagatgt	catccgtggg	tctttaagcc	tcaccaacct	ttcgtcttcc	780
atggctggag	tctatgtctg	caaggcccac	aatgaggtgg	gcactgccaa	tgtaatgtga	840
cgctgggaagt	gagcacaggg	cctggagctg	cagtgggtgc	tggagctgtt	gtgggtaccc	900
tggttggact	ggggttgcgt	gctgggctgg	tcctcttgta	ccaccgccc	ggcaaggccc	960
tggagggagcc	agccaatgat	atcaaggagg	atgccattgc	tccccggacc	ctgcccctggc	1020
ccaagagctc	agacacaatc	tccaagaatg	ggaccctttc	ctctgtcacc	tcgcacagag	1080
ccctccggcc	accccatggc	cctcccaggc	ctggtgcatt	gacccccacg	ccagctctct	1140
ccagccaggc	cctgcctcca	ccaagactgc	ccacgacaga	tggggcccac	cctcaaccaa	1200
tatcccccat	ccctgggtggg	gtttcttctc	ctggcttgag	ccgcattggg	gctgtgctg	1260
tgatgggtg	ggccagagt	caagctgggt	ctctggtatg	atgaccccc	cactcattgg	1320
ctaaaggatt	tgggtgtctc	ccttctctata	rgggtcacct	ctagcacaga	ggcctgagtc	1380
atgggaaaga	gtcacactcc	tgacccttag	tactctgccc	ccacctctct	ttactgtggg	1440
aaaaccatct	cagtaagacc	taagtgtcca	ggagacagaa	ggagaagagg	aagtggatct	1500
ggaattggga	ggagcctcca	cccaccctgc	actcctcctt	atgaagccag	ctgctgaaat	1560
tagctactca	ccaagagtga	ggggcagaga	cttccagtca	ctgagtctcc	caggccccct	1620
tgatctgtac	cccacccta	tctaaccacca	cccttggttc	ccactccagc	tcctctgtatt	1680
gatataacct	gtcaggctgg	cttggttagg	ttttactggg	gcagaggata	gggaatctct	1740
tattaaaact	aacatgaaat	atgtgttgtt	ttcatttgca	aatttaaata	aagatacata	1800
atgtttgtat	garaaaaaaa	aaaaaaaaaa	aaaaagggcg	gccgctctag	aggatccctc	1860
gaggggcccc	agcttacgag	tgcatgagac	gtcatagctc	tctccctata	gtgagtcgta	1920
ttataagcta	gg					1932

<210> 40

<211> 1430

<212> DNA

<213> Homo sapiens

<400> 40

aatttgaccc	tacttccctc	tcagtcctaa	gggcctatct	ttcatcacta	ggttgaatta	60
tctcccatgt	ttgatttgcc	tctatctccc	tatgggcttg	caacaccatg	acgggcacat	120
tgcaagtgc	cttakacaaa	tgagatcaga	tgacctgggg	aacgtggctt	gtacacacct	180
ttctgtgttc	tgtagcatca	gctaagacct	taaaatcagt	aagaaagtat	ctgtctctct	240
gttcacccat	aggaagcagc	ttcgtgggtga	gtgaagggag	ctacctggac	atctccgact	300
gggttaaaccc	ggccaagctt	tcctgttatt	accagatcaa	tgccacctcc	ccatgggtga	360
gggacctctg	tggacaaagg	wcgacagatg	cctgtgagca	gctctgcgac	ccagaaaccg	420
gtgagccatg	ggagccggga	tggggataga	aggtggggaga	ggctgggttg	aaagaggcat	480
tgtgtctcct	ctacctaaag	aacctgggrk	tctgagccat	tgacaagtgg	ctgaataaga	540
aggcccatca	atctaataaa	cactaatgta	tgtgtgcca	ttgcctttca	aggggggaaa	600
ttcttagaga	gccacagact	ctcagagtaa	gaaaggacca	cagagaacat	ctggcctagc	660
tccagacaag	caaaatgct	gcacttcaga	tatccctgca	ttcagagcct	atcttcttgg	720
tactagctgg	gtgatcttga	gcaagacact	gcttaccttt	cagtacaaga	gaatgaaaat	780
agcaccaccc	cacaaaactg	tcattaggtat	tgaatgagaa	gctgtgtgga	aatctcacag	840
catattcatt	aattcactca	acaggtattt	cttttagtagc	caggcatatt	tttaggtatt	900
gggaagacag	cagtgatcaa	aatatgcaaa	atctccaccc	tcataagct	tacatgctag	960
tggggtagac	actaaacatg	cactgtggaa	tatggtagcc	actagctaca	tgtggcattt	1020
tattttaaat	taattgaaat	taaataaaag	taaacattca	ttttccagct	cataccaccc	1080
agattttcaag	tggttccatag	ccacacacta	gcagctacat	tggtgcacaa	catagmtata	1140
gaatatcttc	atcactctga	aaattttctca	tgggacagtg	ctgcagtggt	caaacaagca	1200
ggtaaatat	atgactctgt	taggtgatga	tagatgccgg	tgtaggggaa	aagaatgatg	1260
tacaaagcat	gtggagtgtc	aagctggggag	tttgggtgga	gtttcattat	acagaaagtg	1320
gtcagggggag	gcctctctga	ggtaagagga	tcacttgagc	ctaggagttc	gagtcacagc	1380
tggacaacat	agcaagatct	catctctaaa	aaaaaaaaaa	aagggcgggc		1430

<210> 41
 <211> 1407
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (353)
 <223> n equals a,t,g, or c

<400> 41
 gcttaagatg aaaagttcct tttcttgtgt taatggatgg cacaactggc ataaaaggtc 60
 attaaatgct aatagaccca cttgagggtat gctcgcttaa tggaggatta gagcaaaaca 120
 gacttaaaaag accaaccatgc cagttgtgcc atcccttaag atgaaaagtt ctttttcttg 180
 tgtaaatgta caaagctttt cttttggcac tgacaactgt gttctacctg ggaattttga 240
 atagccattt tcatggctgt gtgttgtgta acacaaatgt ttttaaatgg tattctcacc 300
 cagtaggcca gctctccaaa cgttgcttag atgcttcaaa attagcatat ttnaagttta 360
 ccagtataaa ataccaatgc aactactcta catagccaaa tgtttgtaaa tcacgtctta 420
 ttttctctgk gtttttctact ccaccaaatc ttacaaatsr ttgaaagaaa tatattctaa 480
 cagtacgcac tgaatagtga aaataattag acattttaag aaccagagcc atagaattat 540
 tttaaattag tagaaaagag gagctatttc cgaatctata gaataaagta ccacctaaaa 600
 ctgaatttta tcatataasc aagtaatacc tattagtcac acctaaattt ttcagcactt 660
 cattcaatta aaatmcatga attttaaata ttttacatga tgtgaatagg catgataata 720
 ctttttagtat aaaatctaaa ctttttccat ttatcagaaa tgataaaatc cagttaccac 780
 atatcacggt tataaaatcc ttaattaaat gagtaacttc taaaaataaa caatactaaa 840
 tatcacactg cgatggaggt cccaaatatg tggctcatca ccactgaatt catgtaatag 900
 ataagaaaaa aattagaggt ggatgtcttg ttttgtgtca tgaattacta aaatctctta 960
 gtatgtgtgg tatatttttg agtaaaatta ccatttccag atttgagttt gaagggtctt 1020
 tatagtkgta ttttctctct cactgttaat aatcataatc ctttttctag atttttagtg 1080
 cctgaacaac tggtttatct acaatctcaa atcctaagtg tataattatg tgcattgtca 1140
 atacctcata taatacttgc tcaacagtat agtggtagca tggcattaag atgggtgttt 1200
 tgttctacat atttttcaat atttattctt tctatgttga aattatatca ggctttaccg 1260
 gtttttttag ttgttttaaa aagtaatat ttcaaaagaa taaaataacc aatgatattc 1320
 cttggaataa tctgtaaaac gtagttataa aattctattt tctacttaga aaaaaaaaaa 1380
 aaaaaaaaaa aaaaaaaaaa ggcggcc 1407

<210> 42
 <211> 950
 <212> DNA
 <213> Homo sapiens

<400> 42
 ggcacgaggt taccagcctg ttttaattaca gcagacttcc cacttttctc ccacttagta 60
 tttccaattt gctgttctct gaaacctagg aagaattgaa aattgtctag agaataagca 120
 tggcagattt gttaaatcag cgaccttatt ttatatatat ttctaagtca tggccatggg 180
 catagaagct tcttttttaa ttaagaagga aaaataaaaa tatgtgaaaa gaaagccata 240
 aaggtcattt tacacacatg taactccatg caggaatgcc agtccttccc cttgtgtgtg 300
 cacttgagac tagttctact actatccttc aaaacccaag tgcattgaatt ccatgaagtt 360
 tttcccattt attctcattt taattttcct tctctgaaca actatgacat taatttatta 420
 cttaatcatg aattatggca tacaactccc taattgatgt ttgtgggttt ttttctcccc 480
 cagctagatt ttaatttctt tgaagacaga agccatgctc ttactgtgct agaatatctg 540
 tctcccgtag ctcctgacac agtgctctgt gtagagaggg tgcttgttgg ctcaccaatt 600
 tgttctttac accaaaatgcc cagggaatc ttacatagag ttataccag gcaagaaaag 660
 gatattgctag attctccagc tgccaaagac tggaaatgtca ctggtatcca gtcaccacaa 720
 tctctaggtc cctcattttg ttcttgggtga gaaaggagca ctaaggagat ttcgtccttg 780
 aaaaggcaga aagcaagtg agtatcatct tgccatctag cttggaaatt aacacttgat 840
 cctaaattag gtaatcttcc cttcacatct cagagttttc caggcaacag acactcagta 900
 cgaacaacaa caacaacaac aacaaaaata ccaaaaaaaa aaaaaaaaaa 950

<210> 43
 <211> 1004
 <212> DNA

<213> Homo sapiens

<400> 43

aattcggcac	gagcagcatt	ccacttgcaa	ttggaagccg	agagaaacca	ttgtttatga	60
aakkaaagag	gctkctcaga	tgactgcaaa	ccagccttcc	ttactgggtt	tatcactggg	120
aatgttataa	agacagttgt	ccagtttcat	gaatcttgta	ggtttttgtt	tgtttatttg	180
tttgcttttr	atgttggtgt	tggtgctggt	gttttccaaa	ttcagtattg	tagaaaaata	240
tgctgcccc	gaagagatga	ttggacactc	tccagcggtg	tggtggactt	tgatcatctc	300
tgacacagcca	tctccagacc	ttagtgttta	cctcacgtta	gttttttata	ttctgcaaag	360
acaaamccaa	aataatccaa	atttgacaca	aatacctggg	atacatctta	tttgagatgt	420
ttaacaaatg	tctggatcat	cttttcttac	attggattat	aacgcaggaa	acactgtgaa	480
gtaagttaaag	ttggaattcc	caagtcmag	accatttgaa	tatttacaag	tagatttgag	540
gcaggaataa	tacaggggtg	ccgcagggtg	acaaattcta	ggcagcagat	ttacatgact	600
tgaggctatg	ggctgataag	acgctgaaaa	accaggggtg	ggaccaagct	ggctaagact	660
gactggaccc	aatgtgggtg	tagatttgag	gtagggttta	cctaggccct	cattatacac	720
ttattaacat	actaaatcac	acacccacca	gtgccatgac	agttctgaga	ccaatatgtg	780
atgtaaaaat	ggatggcacc	acagttccga	gaaatcacct	ttaccagga	attttcacga	840
atattccact	ccttggttaa	agaaacccat	tgagatgaaa	ccccagaacc	cattgttctc	900
tctcgggtat	gcccgaactc	ccctttcttg	agtgtgtact	ttctgctttg	caatacatct	960
cttctttcac	tatttgctga	ctcatccttg	acttggttct	cga		1004

<210> 44

<211> 1681

<212> DNA

<213> Homo sapiens

<400> 44

gaattcggca	cgagtcgagt	tttttattcc	tccactgaga	atcacacaaa	aagttagaag	60
cacaaaaagt	atgatgggta	atgatttgct	ccacctcggt	ttcttgcaac	taagttagg	120
tgtagcatca	gggggatgga	ttttgtggcc	actgaggaga	ttgggtgggt	cccatacgag	180
taaggatmca	aataaaaaatg	gmcacsygt	gcattgcttg	gtcattacca	atgagcctct	240
agtttccamc	aagaagattg	ggctctcttc	tcctcacact	tgtccatcaa	ctctccaaca	300
gttttgatcc	ccactgtaat	taaactagta	tcttctaaac	acaaaatctt	cactctacct	360
cagtagcgct	tggcagctga	aatcttttct	atttagaata	tcccaccttt	ctatcttgaa	420
attttgtcca	agctaaatgc	ctctacttaa	tctctgcgta	cctgcgggaa	cacaatgtgg	480
ctaccacatt	ggctaccagg	gctgtaggga	ggattgtctc	aaaatcctct	ccatttatca	540
caraaagggg	ggcgggaara	ggaaraaagt	aggttatgcc	ctgaggctca	aggctactgg	600
atggccaatc	tgtgctaggt	ttgctgggtc	gaaagttaga	tgatatgagc	tgatatagsa	660
gagaaatata	gggtacagtt	tctaccctga	ggggctgtat	tttagttggg	gagatacatg	720
caatgactgg	acaccaccac	caaggataag	gaagtcctgg	gattgtgtga	aagccacagc	780
agttcagaga	ggagaaggaa	aaagactcca	tggaaatgat	gggaattgaa	ccaggcctgg	840
gtttttcccc	tctcaggcac	actggaggct	gtttgcctac	cctgttgcat	ctcttggtct	900
ttccaagttt	ctgtcttggt	acagactctt	tcctctcttc	ctcctcctag	aaatattggc	960
aagcttcttt	agtcatttgt	gtttctttac	attacaggcc	agaggtgtat	cttctctgat	1020
agataatggc	cctcagttaa	gactagggaa	agctattttg	cctgtctgat	tagcgcccta	1080
ttttagaata	atcctattcc	cttgattctt	tagtattttac	aatttttcta	agtaccgatt	1140
atattttcta	agtcaaagtg	gggtaaaatt	agtgcattgt	atcctgttgt	tgccgctttc	1200
tggagtagtc	agtcttacat	atttgaacaa	taccaccctg	gtgtaatttt	aaaaagtaag	1260
agcttgatcc	tttaaaaaac	atttagccag	gcagtgtgag	ctctctctga	ggatcctcac	1320
attaggagtg	ttttacatac	atcacacaaa	aggaaaatgc	gttctgaggg	gatcggggct	1380
cctccgagct	gagagctgga	cctgatgaat	tgtgacaaat	gggcctgttt	ctgccagctg	1440
cacgtttcca	gccaggtgac	gtctgaggct	gcctgccagt	aatggtttgt	ggtttgggga	1500
gcaagagggg	ggccctggac	atactcactg	gtgggggaaca	ggaaaaagtc	aggcccaatc	1560
agaaaatagta	actctcctca	gtgttcccc	gctaagtaag	actatgcatt	taccatacag	1620
tccccatcct	aaaactcatg	aatgaagaa	ttagtgcac	actgggggag	tagtggctcg	1680
a						1681

<210> 45

<211> 1365

<212> DNA

<213> Homo sapiens

<400> 45

ggcacgaggg	agaactgctt	taattagcct	aggtagaaa	tagtcctagc	agtgtaaata	60
tgtataatta	gagttttcta	atttcactgt	gagatctcta	acttttgagt	ggcaaacaga	120
tcaagtcttt	tgctcataga	cttttctgtg	gggttattaa	aatgcaaaag	ctttattttt	180
tttaataatg	ccatactcca	ttagtgtcag	atgatggat	ggaatttggt	cccttgcttt	240
ccccactgt	tactgcttca	gtttatagat	tgccagcaga	gttcagaaat	agagcagggg	300
tttaccggt	ctttgcttgg	acatcccat	ttcttttgct	cagacccatg	ttggcaatca	360
tgtatgaact	gtgttatact	tctcagtgt	ttcttttttc	tttttgataa	gatggatata	420
aaaaatagtt	gctgtgcaaa	agtttagtagt	cttcttcaag	aagaaaacca	attctttttc	480
taataatata	ctgtgaaatt	gcttcattca	ttcatttatt	tttaagccaa	atg.cagcag	540
agtgtctgtg	cttttatcta	gtaattttga	tatgtaagta	ttaatgcatt	tttaaaagat	600
gtctacattg	aaacatgttc	ttcccaggt	cctgcttatg	atgctttggt	cagatttttt	660
gtaagagacc	agtttagtaca	ctgggggtgt	atattgtgta	catgtgtcat	tttagttagg	720
cattgtaggc	caaatgtgat	tataaatgaa	gttgatgaac	attaattttg	ttattagtga	780
gttttttgaa	ttgtaaatgg	atttccagtt	tacctctgt	tgtctacagc	ttttttaatt	840
ttaaggtttg	actaattgta	tccatctcat	tgtacagtgt	tttagttgca	agcagaaagt	900
agaatttggt	ataaagcagg	ttatttctat	attgaaagga	gtacagttga	aattgttagt	960
ttaagattgt	taaaatcatg	acaattctaa	cttgcttatt	ctaacctatt	gtgtacaatc	1020
tgatttttta	aaattgtaaa	catgtatgat	cttggtttca	tgtgtttttg	aaagtgttat	1080
tgtttaaaaa	atgaaaaaag	catatctgct	aaagagctgt	cagttttcat	tactgactct	1140
gtaaaatata	ctgttctttg	tgtactgtgt	gttattttgc	cagctgctgc	attagccttc	1200
aaaagtattt	ggaaacttaa	gatgaactac	atttcttgca	aagtacattc	ctttctgttg	1260
tattttgtcc	tgtaactgaa	gtatagtaat	tattttatgg	aaatgttagc	aattctgtac	1320
caactttgaa	taaaatgaaa	aatttataaa	aaaaaaaaaa	aaaaa		1365

<210> 46

<211> 1137

<212> DNA

<213> Homo sapiens

<400> 46

gggctttttg	tcaacctgaa	gcacgttcta	agtcgatggt	agaaagtgga	cacccccaga	60
agacactttt	gccagaaat	ctctttcttc	ctgacccctc	ttccccagag	tgcccggaat	120
tccactgtca	gaaatgcatt	gtctgggtta	aaaaacttaa	cacctgctat	gatttcaaca	180
gtgtcaaaac	aggatacgtc	aaaactgggc	gaggaggaaa	tgtatttggg	ttctaggata	240
gtgaaaagctc	tattttttct	acttttctgt	attttccata	tttggtacaa	tgagcacgta	300
cttagaacgg	tttttagattt	acgaaaatat	gcaaacacag	tacagatagt	tcttgcgctc	360
cccattgcta	gttctctcat	tgtcaacgtc	tcaacgttag	tgtggtgctg	ttgttgcaat	420
gggtgaatga	atattcgtgg	gctgttatta	aagtcagtg	ttcaccccta	tttccccagc	480
tttctcttta	catccttttc	tgttccaaga	tgcatccagg	atgcgcgctt	acattagtct	540
tcacacttcc	ttaggttcct	cttggtatga	tggtttctca	gatttttctt	gtttttgata	600
atcttgacag	ttcagggagt	atttgcagg	cattttgtca	aatgttcttc	aactgggggtc	660
tctggtggtt	ttctcatgat	tagtctggga	atgtgctttt	gggagggaaga	ccacagagat	720
gatgtgccag	tctcagaaca	tcgtactaag	aaaaggttct	gccaacttga	cttaccactg	780
ttgctggtga	ctttgagccc	cgggctgagg	ttactccttt	gtaaagttag	tctttttttc	840
tcctttccat	gctgtatgtt	ttagaaggaa	gtcactatgc	tgtccaagc	aactcaagtt	900
tgatgaatgg	ggagttccgc	cccacctcct	tgagggcaga	gtagctacat	aaattacttg	960
gaattttctca	aggagatttg	tctgtactcc	cagtttatta	tataaataaa	tgatttattt	1020
atattacagg	gacccagggg	tctttactgt	atgctttggg	ttataattca	atggtacttt	1080
actttgtggc	tcaagtatac	tactttttaa	ttggaaaaaa	aaaaaaaaaa	aactcga	1137

<210> 47

<211> 2763

<212> DNA

<213> Homo sapiens

<400> 47

agagtttgac	cctggaaagg	tgctttgtat	atgttctttt	cacatagtgc	ccagcttgca	60
tgaaatgtac	agagaaatgt	gtggctgtat	tttttacttt	tgtcttgat	atgtatgtat	120
attgggtcct	ctgggcagta	gaggcaaaag	tcacctocca	tgtagcacat	gaaatgcttg	180
tgagttgttg	acattggaca	ggtgaacagt	agggcattac	atttgtgtga	attaaatgtg	240
aacttctgta	ttacgttgcg	gcgtcggcag	tcctgcgttc	cctggagtaa	ctgtacgtat	300

```

ctgcctttgc tgggaagact gyggggctgc ctgtgttggc tggcgaccag caggattgct 360
ccaggatttt gtgtttacct cgcgtgaagt tcagcacgtg ctgtcgtgta gtcagcttct 420
actctaattt ctgttacagt tctgcaaagg taacctggag tttagaagt aaaaaaagc 480
atgggatgtt ggatttgac catttggagt ttctttaggg aagaaaagt ttctgcttt 540
ttatagaaaa tcatttcagt ctcccgaggt ctcatgctag caaattttga aataggattc 600
taatcactga tttcaaatat taagcaaat gtaaagcact ttaatttata gctatggtta 660
taaacagggt ttagatgttt caaatgactt gtccactgaa tgtcacttga ccttgataag 720
aggcgccttg cacacagagc ccagttaatt ctccgcacct cggttgtgtg cttccgaatg 780
ggctcactcc cgtggtgggtg tttgagagcc aacaacacta cctcagagac gggctctttg 840
gaaacttttg gtctcactgt tgc-tggctg gaggactttg gtttatagct ggaatactga 900
gtttcagttca gaaggcagga aagacagtca caccgacgtg tctgaagggt gtaggctctc 960
cacttagtgc cacaagctga cggctgcagc cagcaggccc cggtgacgag acacttccag 1020
gtcttgtggt ggggacgcct ctcatgcca gtcccgccac tgcctgagtg gcctgggtgt 1080
cttgcccttct tggaaattac tgctcacctg gtatctgtac gttaatgttt cttgtctgag 1140
tacagttttg ataaagaggg tctcatttcc tgtgtcttgt atattcagtc ctttcaatac 1200
gtccaccttg aggcctcacca cttggagaga cacagggaag taatatttac agctgtcatg 1260
tgacatcccc aggtctttgt gttttgccct gttttacggt gaggtaggag ggaacccatc 1320
tggggacccg taggtgcagg tgcagtagga cgtgggactt ttggacccgt cctttggtgc 1380
agctcgccag ggatgagagg cacctcccta cttgggtctt caggagctgg tccaaggagc 1440
ttcgaatcta agtcatctag aatgacctg aaatgactga cagcccccgg cccaagaaaa 1500
acccataacc acctcagatg gatctgacgt ggctaaggga caaacagcaa atatttcagt 1560
catttttgatt ttacaaataa aaaatgtgtt gtgtttttgt ccgacattat ttcctgactg 1620
cactgttctg agaatggagt ccacctggtc cctctggttg attagaatct caggtttcag 1680
ctcctgctgt cctgagcgaa cttgcctgat gcagggtgtg gctgtgtcca gatgttgctg 1740
gggctcact tttctcttg gctggaggtc caattgccag agcctccac actgcacata 1800
caaaaggttg agccaggggc agcttctggg gccactgcac agggccacct cttgggttcc 1860
tcggagttta atttgaaagt ctgggtgtct taggatgat gttaggaaac ttgaaaaatg 1920
gctgcaaata gccaaatcaa acttaagaac cagatctctg ccagattaaa cttttttgaa 1980
gcttttaaaa gtcaatattc ctagtggcca ctgagttcca ggcacactgg tgccctttac 2040
tgccacagct gctcaccttg tctggcaaac tggagggacc tcagaaactg gactcctgca 2100
tgtccttggg ggcgcagccc tgtgtgtctc aggcagagct ctcaggagcc ggggcacctt 2160
gctgttctgt gctgtgtcgt cttctaagt gagctcatcc actgctgctg cagcgtgggtg 2220
atcaggagtc acagacaaga tgggggatgg tgtgtgtgtg tgtgtgtgtg tgtgtgcacg 2280
tgtgtgtggc taaattaagt catactgtca accacacgtg atctcgtctg aaacagtgtt 2340
tggaaagtggg aacagttttg tctgtatgc tgatgtgtcc agaatttcat ttaatgatag 2400
acggaaaaatg tgtggttact gaaaactgta tatgatacag aatttcataa gagccatgct 2460
gttgggcaaa gcaactcttt ttcaaccact gctcatcagt ttctgtagag acaaaaaact 2520
tgtacatatt ttggaatctg aagaatccta tgtaaatcat ttgttactta agtctgtgaa 2580
aaacatatatt ctttggagga aaatgtatgc atttataagt gtcccatgga atcagttttt 2640
attgtatcga tataattgtc tctaagtgtt gactgtcttc attgcaatat gaaattcatt 2700
aaaatgtcca tgttccataa ttactattat aaaaaaaaaa aaaaaaaaaa aaagggcggc 2760
cgc

```

<210> 48
 <211> 1576
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (252)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (804)
 <223> n equals a,t,g, or c

```

<400> 48
ctgctgctgt gtccctgggtg gctgtgtttt gattgggtcaa tgggctgcat cccctcatt 60
aagtcacatca gcgactggag ggtaattgca cttgcagcac tctggttctg cctaattggc 120
ctgatatgcc aagccctgtg ctctgaagac ggccacaaga gaaggatcct tactctgggc 180
ctgggatttc tegtattccc atttctcccc gcgagtaacc tgttcttccg agtgggcttc 240

```

```

gtgggtcgcgg antgttcctc tacctcccca gcattgggta ctgtgtgctg ctgacttttg 300
gattcggagc cctgagcaaa cataccaaga aaaagaaact cattgccgct gtcgtgctgg 360
gaatccttatt catcaacacg ctgagatgtg tgctgcgac ggcgagtggc ggagtggagga 420
acagcttttc agaagtgtc tgtctgtgtg tcccccaat gytgaaggttc actmcamcat 480
tggcaaaaac ctggctgata aaggcaacca gacagctgcc atcagatact accgggaagc 540
tgtaagatta aatcccaagt atgttcatgc catgaataat cttggaaata tcttaaaaga 600
aaggaatgag ctacaggaag ctgaggagct gctgtctttg gctgttcaaa tacagccaga 660
ctttgccgct gcgtggatga atctaggcat agtgcagaat agcctgaaaa cggtttgaag 720
cagcagagca aaakttaccg gacagcaatt aaacacagaa gggaaatacc cmgactgtta 780
ctacaacctc gggcgctctg taancgcggg gtgccctgtg cctgtggaag gaaagatggg 840
ttatttttct tatttataat aaaatgacat agtgacaccc acctagccca tacattttat 900
aaagttcytt cacatgtttc tayctcattt gaaggtagct atttgatttc cttttgagta 960
attttttaaa gctctcatta gagagcagta cagtgtgaat tagtcaagtt taagagggtca 1020
cccacgcaaa aggttaaaac caggaataaa ttaacatgtt aaagteccgt ccgccctgta 1080
aaacagcact ccaatgggta acttctgat aaacatcagt ttctctgttt ttaaaacaag 1140
aattgagtaa gaacagagat taaagtaaca aatycgtagt atgattttcg agctcccttg 1200
ttctccttct tcaagggagc agagctcttc atctgcaggg agcatttccc ccaaaaaagg 1260
cagctttgga gggcacggga tttatttgaa agggctttga cattatttgg tggaaataga 1320
aaataacgtg ttctgtagta gctttatatt tttggttatt gacaggatgt ttacgaagat 1380
ctgattgtct ttgattttct tgacaaaaat aaaatgagac acacacatag caaaattctt 1440
taaacacgaa tggttgtctt ctccctataa tcaatcattt aatttggttt caagaaaaca 1500
aatacatatg ttcctaarat atttagatgt attcaataaa cattgttaat taaaaaaaaa 1560
aaaaaaaaa ctcgag 1576

```

<210> 49
 <211> 1348
 <212> DNA
 <213> Homo sapiens

```

<400> 49
gttaaaaaac atgtaaaacc gtatttatct tcgaattaca gtgttatgtg tttgaatggt 60
tttagatgtt aaaaagtagc aaattgaac ttaatgttta aagtctttgt taattgaaaa 120
attgatcttc aatagtggta ctatttgcag tatgattcgt tcttttaagt tacatacgta 180
tatattagta catacgagag tgatgttaga cctgtagaaa tgaagggtgt gttttaattg 240
aaaacattta tgtttatttt gctgatagtg tttgtatttt caaaaagtaa acaagttctg 300
tcaatatgtt tgaaaatttt taaagttgag ataaatagca tctcattttg taaaaataaa 360
aaatataaag atttaccata tgcgtttgca tcagaaaaga ctggaaggac atactcaaat 420
gtcaacaatg attatctctg aatatgggat tatgggcaga tttttatatt ctttttactt 480
atctgtattt tcaaaaaact ctacagtaag tgaactgcat ttataatact gttttaaaag 540
attgaaccac caaagataga ggttattaaa aatttatccc ctactcact gattatagta 600
attggattat ttttggattt caagaaacat tagtattagt ttaagagaat gttgctatat 660
gtaaagcatt gtactaaaaa ctatgggaga tatacagaag gaaaagatag cttactttca 720
aggaagctgt atttcaaaaa atgtgtgtag aaagtgccag agtggcaagg aaatttgctc 780
accagttatc ccactcctta atacagtttc ctggcaaatc tttgtttctt tcttagacta 840
atacttgag acctatgtct ccttgactc ttctttcaaa tctaactttg tttttttaat 900
ggatcatgaa agataaattt ctgtaattga tgttttatcc atagcatgaa gattttcttc 960
taaactgttt cttccttttc tggtaatcat ttacagtggg ctttatgtta caatttgaaa 1020
cacagtagaa gtacaaaaat atggccaggc gcggcggctc acgcctataa tcccagcact 1080
ttgggaggcc aacgtgggtg gatcacttga gctggggagt tcaagaccag cttgggtcaac 1140
atgggtgaaac cctgtctcta ctaaaaatag gaaaattagt cgggcgtggg ggcacatgcc 1200
tgtaatccca gctgcttggg aggcctgagg acgagaatcg cttgaacatg ggaggtggag 1260
gttgacagtga gccaaaggtg caccactgca ctccagccta ggcaaccaag cgagactttg 1320
tctcaaaaaa aaaaaaaaaa aaaaaaaaaa 1348

```

<210> 50
 <211> 1264
 <212> DNA
 <213> Homo sapiens

```

<400> 50
gacccacgcg tccgcccacg cgctccgttt cattcacatt cacaagcaa acatctagta 60
catgtctttc acttcacttt atgatagtg attggatgat ttgggcatta cgatcacctc 120

```

ttaccacagc	acagaacata	cattcttcaa	cagcattaac	ggagtttgcc	aagtgcatta	180
aagaggtcac	gtggagggtg	cggtcatatg	aaacaatctg	cagaaagtgg	ggtaagaaag	240
ggcacatggc	acagttaaag	ttgtagaaat	caaattacta	tcattttttg	ttgccaaaac	300
aaagtcttac	atttaacccc	cctttctacc	acccccctcc	acacttcacg	tcagctacat	360
agtttccaca	gggtaattca	ctaagagctt	gtggagcttg	gtttttaa	ccttagcctg	420
gtctgacttt	aggcatagct	tcagttcttc	ttccgtgtcc	tgggttcttg	ttcagtttta	480
cttctaatcc	accaccaaaa	gaaatgtctg	gctggtctca	gctagagtct	atgtgtctta	540
gagcatgtgt	gcgtatctga	ccatcatccc	tgctctcatc	tcagctccct	ccaggctgag	600
caccggttcc	ttttgtccca	tacgtcatga	agtcactat	tgggaaacct	gtgcttcctt	660
ctccatggct	taactccctg	tcagtgtcgg	agtgataaag	aatgcttgta	aatactgtaa	720
tatatattatt	aattattgaa	aggcattcat	tcagtggaca	gtgggaatta	actctcccaa	780
ggcaagtga	aatgaatgat	tgacgtacgt	tgatttaaca	atcttactag	attttaattc	840
ttaaggattt	caaatgaaac	cagaagggtg	ttatgtaaga	ggcttaaaat	gatcttatgt	900
ttaaagagat	tctgttatta	gcaccatgaa	ctcgtactat	gaaattttta	agccttttat	960
ttttctaact	atattactgt	aggactggat	attaggtgtc	atataggaag	cacaaaagtt	1020
tattgtctgt	tgctaaagca	aaatagcaga	aaattttgta	tatgcaaaac	tggtgaagga	1080
ccatagagaa	tgtgtactac	tgccgggctt	ttactaggct	tcctgcgtgt	gtaaaagtcg	1140
aggattgtct	ggcattcagg	gtgacatgat	ggactaaaat	gttttccatt	aaagtcttct	1200
attttaaaat	ttagagaaaa	ataaaatggc	tttccatcag	aaaaaaaaaa	aaaaaaaaaa	1260
aaaa						1264

<210> 51

<211> 1660

<212> DNA

<213> Homo sapiens

<400> 51

accacagcgt	ccgtatacat	atctattagt	atagtatctc	ttgaatgcga	ttttttctag	60
aatgtgttct	gctgatttgt	tttagagcca	tgagtgcaat	ttatacacat	acatctattg	120
ggaatgctca	gaagtgtgtt	actgatggaa	gtgccttcag	aagagtccgg	gaaccacttc	180
ctaaggaagg	aaagagctgg	ccacagttag	agcaagcctg	cctggggccc	tgctctgtgt	240
tccagctgca	aactgcctgc	atcatccctt	cctgttactc	ttccttcacc	tgagacagtc	300
gaggccacag	cgtcagccag	ggccagagct	gggatttgaa	cccaggcact	cgggctccag	360
agccacactg	cccagtggtg	gggtttagt	gcgctcctg	gccctgactg	aggggctgac	420
tgaagcctgg	tgagagcgtg	ctgggtcagc	ctctccctgg	cgggaatcct	ctccgtccag	480
tcttctaacc	tagcagcctc	acgtccacag	agctgccttg	tgaactcag	cagagccctg	540
gcttcctgca	gagccgtgtt	ctcccagcct	gcttcatggc	tccttggttg	agccaagctt	600
gcggtatccgt	gggtgaaagg	taccgcacc	gcctgggctt	tagtggtatg	tacgggcctg	660
catcgtgagc	agcgggcggg	ggcccaggca	ggtgaggcag	ctggccacaa	gggcagggcc	720
cggcccccct	cccaaggctg	tgtctsatat	tcttgagcct	gttcgagttt	ccttttccaa	780
gectcctgtt	ctcccacccc	camccctgcc	atgctgcagt	gactaaatct	gtggttctca	840
tccttggaag	acacctgagt	agctgctaca	agctggccac	agcccagcga	ctctgatgtg	900
gttgggctgg	gtgtggccag	cagccagggc	atcgggactt	ttcgaagctc	ccaggtgact	960
caccggcagc	tggggatgag	aactgtcagg	agggaaggct	agaagtccca	ggatgcactt	1020
gaaaagcctc	tagctccacc	agtgaccagc	tcctggctgg	actcctggtc	tggaactcag	1080
atcagggagg	ctctggcctc	tcgccctcag	gctgggggct	tcctcacatg	gtcatcaaag	1140
acttggccag	ttccgcctct	cccacggcgg	tcctgtcttc	acccagcatc	acgacgcac	1200
agttcaccaa	caaacacgat	tcagtgtctg	tagtgctggg	tcctgttctg	ggggctgggtg	1260
atgaggccaa	gagggaaaga	gggagctctg	tggtccatcg	agggggcgac	aagcctggac	1320
cagatgaaag	tgactcatgt	tgtaatttag	cggcttaggg	ccaaagggtg	cccctggacc	1380
agtggcctca	gcatcacctg	ggaaaagggt	agaaaatgaac	attcccaggc	cccacctcag	1440
cctcctgaat	cagagcatcc	ctttggcaaa	ccatactgag	aaacaaccac	gtgcatcacc	1500
aagcgctgtg	aagaaggcaa	gctttgagac	cttgaaggaa	tcatacaact	ctgggcctcg	1560
gtgtgctcac	ccagggcgag	acaaagacgc	catgctcttc	caagtggcct	ccaagattaa	1620
atgagcaatg	acttttataa	aaaaaaaaaa	aaaactcgag			1660

<210> 52

<211> 1678

<212> DNA

<213> Homo sapiens

<400> 52

aattcggcac	gagccaagct	gcactattgg	gaatggattg	tggctgaaca	gcaaatcaaa	60
acaccagaaa	tattttttata	tgttaacgct	atattatggt	aatggtgctg	aaaacaaaac	120
ctaacaacc	ttgatgtacc	agtcgaatac	catgtagcgc	tgagtataaa	agttaaaatg	180
tgctgtgctt	cccacccttg	tcagagggaa	gggtggctat	gtgttatttt	cactgtcttt	240
ttgaaagtta	cagtattgtg	tttcactttc	gtgcagataa	ctggaagtaa	agcggcaaac	300
agtgtatta	catgctaaag	ttaccttctc	tttgtttttt	gcatactctg	aattacacct	360
ttaaagactg	atatgaatca	gtacgggtcac	tatacatctt	atgatttttc	tgcatcttta	420
aaattgtatg	atcgtaacat	tatttattac	cacaaaacag	caaatctctc	aatgtctaag	480
aaaactagct	taaaatgttt	aaatatagtt	ctgattgggt	attaattact	tgattaagaa	540
aaaattaaaca	ttatagatac	tctggcatta	cgcttctata	ccttttaggt	cttctctgca	600
atactggaac	ataattcttt	tgtgtagctc	actattagcc	agctaagttc	atctttttaa	660
taccataaaa	aggttatatg	tacagttcct	attttagctt	gcttacaaag	ggagcattat	720
ttttatttaa	agtattgcta	gtaaatgatt	tgtagaaact	tgggttttcta	agcatagttc	780
ttccataaac	accttttggt	gtttgagcac	aagggattct	tttcttagtt	ctatgtgttt	840
gtttccctat	atgcagtcct	taaaggatta	caacacttaa	aattgaaatg	actgtgtgca	900
agctttttgc	atcatacatt	ttttgaaaga	tttttaaaaa	agcctacaac	ttacatatgt	960
agtagaatca	gccattgctc	tgtctctggc	atagagtcac	ctgttatgtg	gattaaaatg	1020
ttttaaaata	catatttgaa	gmcctttgag	aatgctttag	tgtttgattt	gaaataaaag	1080
gaaatttttag	caaggattaa	agaaaaaagc	tatcagctgt	atgttaagag	agactcttac	1140
taacatgttg	taaatattac	aattcatgaa	atgttattgt	aagtctgtaa	cttaattttt	1200
tcctgtttt	agttatcacg	gttgggttgg	aaatttgggt	tttggcataa	acaagtaaaa	1260
tgtgcccatt	ttatggkttc	catgcttttg	taatcctaaa	aatattaatg	tctagtgtgt	1320
ctatattata	accacatttg	cgctctatgc	aagcccttgg	aacagaacat	actcatcttc	1380
atgtaggacc	tatgaaaatt	gtctattttt	atctatatat	ttaaagtttt	ctaaaaatga	1440
taaaagggtta	ttacgaattt	tgtgtacaaa	aatctgtaca	aaaatctgtt	tttacatcat	1500
aatgcaagaa	ttggaaattt	ttctatggtg	gcctagttaa	ttgagcctgg	tttcaatgtg	1560
agaaccacgt	ttactgttat	tgtatttaat	tttcttttcc	ttttcaacaa	tctcctaata	1620
aaactgtctg	aatctccctt	gtgactttta	aaaaaaaaaa	aaaaaaaaaa	aactcgag	1678

<210> 53

<211> 1860

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (912)

<223> n equals a,t,g, or c

<400> 53

cctagctgtc	ccccctgagat	gaagaaagag	ctccctgttg	acagctgcct	gccccgctca	60
ctcgagcttc	accctcagaa	gatggatccc	aagagacagc	acattcagct	cctgagcagc	120
ctgactgagt	gcctgacggg	ggaccccctc	agtgccagcg	tctggaggca	gctgtacctt	180
aagcacctgt	cacagtccag	ccttctgtctg	kagcacttgc	tcagctcctg	ggagcagatt	240
cccaagaagg	tacagaagtc	tttgcaagaa	accattcagt	ccctcaagct	taccaaccag	300
gagctgctga	ggaagggtag	cagtaacaac	caggatgtcg	tcacctgtga	catggcctgc	360
aagggcctgt	tgacagaggt	tcagggctct	cggtgcctct	ggacgcggct	cctcctgttg	420
ctgctggtct	tcgctgtagg	cttctctgtc	catgacctgc	cggtcacaca	gctccttcca	480
ggctggctgg	gggagacact	gccgctctgg	ggctcccacc	tgctcaccgt	ggtgaggccc	540
agcttgacgc	tggcctgggc	tcacaccaat	gccacagtca	gcttctcttc	tgcccactgt	600
gcctctcacc	ttgcgtgggt	tggtgacagt	ctcaccagtc	tctctcagag	gctacagatc	660
cagctccccg	attccgtgaa	tcagctactc	cgctatctga	gagagctgcc	cctgcttttc	720
caccagaatg	tgctgtgccc	actgtggcac	ctcttgcttg	aggccctggc	ctgggcccag	780
gagcactgcc	atgaggcatg	cagagggtgag	gtgacctggg	actgcatgaa	gacacagctc	840
agtgaggctg	tccactggac	ctggctttgc	tacaggacat	tacagtggct	ttcttggaat	900
gggcacttgc	cntgatattc	cagcagtagg	ccctgccttc	ctggccactg	attctctgat	960
gggtagacca	tccaagactg	cagcgggttag	aaggtggcag	ttcttcatgg	gagtcttttt	1020
aacttggtgc	ctgagttctc	tcctaggcaa	gtggccagtt	gcctccacct	cagttctctc	1080
atcttttggtg	gggacagggc	ccagcagcat	ctcagcctcc	taccacaat	tccactgaac	1140
actttttctg	ccctactgca	catggccccc	agcctccatc	cttgtgctgg	tagcctctca	1200
caactccgcc	cttgccctct	gccttccact	tccttccatc	tcatttctaa	accccaacaa	1260
gctcatctct	aaaaagatag	aactcccagc	aggtggcttc	tgtgttcttc	tgacaaatga	1320
ttctctgttc	tccagacttt	agcagcctcc	tggtccatt	cttgggtcaca	gctctagcca	1380

cagcagaagg	aaaggggctt	ccagaagaat	atagcaccgc	attgggaaac	agcagcctca	1440
cctccacctg	aagcctgggt	gtggctgtca	gtggacatgg	ggagctggat	ggaaatgcct	1500
ctcacttcaa	aatgccccagc	ctgccccaaa	tgccctctaag	cccctccctg	tcccctccct	1560
tgtagtcccta	cttcttccaa	ctttccattc	cccacatgc	tgggggtctt	ggtcacaagg	1620
ctcagcttct	ctccactgtc	catccctcct	atcatctgta	gagcagagca	caggcagttg	1680
tgtgccttgg	gcccagggaa	ccctccatca	acctgagaca	ggactcagta	tatggttctt	1740
gggtatgccc	taccaggtgg	aataaaggac	acagatttga	tttctaraaa	aaaaaaaaaa	1800
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1860

<210> 54

<211> 1663

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (975)

<223> n equals a.t.g. or c

<400> 54

aattcggcag	agttttctga	tcagactctt	tttattgttt	tgttttttat	aaacaagtct	60
caggtggaaa	aagaaagaaa	gggaggagct	agctctctgc	cttctcagcc	aattgaaatc	120
gtggaaccca	atgggcttca	gctagcccca	ctcatcactg	ctggggggga	aaagacatcc	180
ctactccctt	tccccgtggc	actcatgata	ttctcaatgc	cccaacaagg	gtcatcttgg	240
ttcctctcgg	cgtttctgtc	ctggcctttg	gctctggctc	cggctctgac	tccgactccg	300
gctccggcca	gggccccggg	agccccatga	gctgctggag	cccctggaag	agttgctgcc	360
ggccctggaa	catgtgctgg	tgccctggcc	ccgggacagg	aagcttggtc	tgctgtatgg	420
gagccaggcc	tcttcatctg	ggtggagcac	ccgctgggct	gccaggggca	cggcctggac	480
cgctttcctc	tccccactgc	gctcccgcgc	cagggaggac	atgctgcctg	ctgccctcag	540
ctctagggcc	cagctcgcct	cttctctctg	cggaggcaag	ggtgggtggg	gcaagtcccc	600
aggactgttc	tccctcctgt	agggaagagc	cttgggttcc	ttccggaatc	gagcacgggg	660
tccttgaagt	gggggagtc	tctccccatt	cccctgccag	gttctgcctg	gggactgct	720
ggctgtgcta	ggggcaggac	tggggctgag	gtgggggtgag	gctgcagggc	cagcacccaa	780
gccagcaggc	ctcgcttcac	ggatgccag	catgggctgg	gatacactga	gaggggaact	840
cggcccaagg	ggcaccytc	tgggcatctg	atggagatgg	ggcatgtcag	tgggggctg	900
gggagggtca	ggaggtgagg	gtgtaagagt	ggctgtggac	tgctgtccat	aggaaggtgt	960
gggagagggg	gtttnccttc	gggatgggg	gaccaggcac	cctccactgg	agctgggctc	1020
cgtcagggtg	cttctctcag	gcattggcgg	gcaccactcc	tctggctctg	agctgccctc	1080
cagctcctcc	tccggccctt	ctaggcagct	cagttcaca	gaaggaggag	gtgggggcag	1140
ggcttctggc	cagttcagag	agggcatctg	cacaggttcc	cccagaagct	tcactttgcc	1200
tcccttggtt	ccactgtccc	cctggctcca	ctctggagga	gcgtactggc	tccagggacc	1260
cagatctcct	gagggatgtt	gggggaagcc	cccatggaa	gtctgcagct	cctccccgcg	1320
tgggtcaatg	gtgctataga	caggaccctc	gccaggggcg	gccgtgcccc	tggccgtctg	1380
agctagatac	agggagattc	ccgcttcgtt	gtaatatctg	tgcgtccggg	caggattgct	1440
agggcagcag	cttccccctg	gttctctggc	cagggggcct	cgagatgggt	ggggccacga	1500
atctgccagc	catgagtagg	gggctttccg	tcctcgaaact	tgccccctct	ttatggagat	1560
ggttgcaaa	cctggcctcc	tcgtggcgct	ttagaggcaa	acgtcatcca	gatccccgcc	1620
cgtcttggcc	cgcagccctc	cctagtctctg	gcagctcctc	gag		1663

<210> 55

<211> 1632

<212> DNA

<213> Homo sapiens

<400> 55

ccgcgccgcg	ggcgcatgtg	gggactctgc	ggcttctcag	gtgggtggag	aaaaggcgct	60
cggtcatggg	gatccagacg	agccccgtcc	tgctggcctc	cctgggggtg	gggctgggtc	120
ctctgctcgg	cctgggtgtg	ggctcctact	tggctcggag	gtcccgcggg	cctcagggtc	180
ctctctctga	ccccaatgaa	aagtacctgc	tacgactgct	agacaagacg	actgtgagcc	240
accacactct	ggggctgcct	gtgggcaaac	atatctacct	ctccaccoga	attgatggca	300
gcctgggtcat	caggccatac	actcctgtca	ccagtgtatg	ggatcaaggc	tatgtggatc	360
ttgtcatcaa	ggtctacctg	aagggtgtgc	accccaaat	tcctgaggga	gggaagatgt	420

ctcagtagct	ggatagcctg	aagggtgggg	atgtgtgga	gtttcggggg	ccaagcgggt	480
tgctcactta	cactggaaaa	gggcatttta	acattcagcc	caacaagaaa	tctccaccag	540
aaccccgagt	ggcgaagaaa	ctgggaatga	ttgccggcgg	gacaggaatc	accccaatgc	600
tacagctgat	ccgggccatc	ctgaaagtcc	ctgaagatcc	aacccagtcg	tttctgcttt	660
ttgccaaaca	gacagaaaag	gatatcatct	tgccggagga	cttagaggaa	ctgcaggccc	720
gctatcccaa	tcgctttaag	ctctgggtca	ctctggatca	tccccaaaa	gattgggcct	780
acagcaaggg	ctttgtgact	gccgacatga	tccgggaaca	cctgcccgct	ccaggggatg	840
atgtgctggt	actgctttgt	gggccacccc	caatggtgca	gctggcctgc	catcccaact	900
tggacaaact	gggctactca	caaaagatgc	gattcaccta	ctgagcatcc	tccagcttcc	960
ctgggtgctgt	tcgctgcagt	tg+cccccat	cagtaactca	gcactataag	ccttagattc	1020
ctttcctcag	agtttcaggt	tttttcagtt	acatctagag	ctgaaatctg	gatagtacct	1080
gcaggaaaca	tattcctgta	gccatggaag	agggccaagg	ctcagtcact	ccttggtatg	1140
cctcctaaat	ctccccgtgg	caacaggctc	aggagaggcc	catggagcag	tctcttccat	1200
ggagtaagaa	ggaaggagc	atgtacgctt	ggccaagat	tggtcagttc	cttgatagca	1260
tcttactctc	accttctttg	tgtctgtgat	gaaaggaaca	gtctgtgcaa	tgggttttac	1320
ttaaaactta	ctgttcaacc	tatgagcaaa	tctgtatgtg	tgaataaag	ttagcatag	1380
catacttcca	gaggtggtct	tatggagatg	gcaagaaagg	aggaaatgat	tcttccagat	1440
ctcaaggag	tctgaaatat	catacttctg	tgtgtgtctc	tctcagcccc	tggccaggct	1500
agagggaac	agctactgat	aatcgaaaac	tgtgtttgt	ggcaggaaac	cctggctgtg	1560
caatacaawr	kgctagggcc	cctgtgtgat	attgaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1620
aaaaaactcg	ag					1632

<210> 56

<211> 2233

<212> DNA

<213> Homo sapiens

<400> 56

ggcacgagct	tgatttgata	tggtaagcag	taatatTTaa	aatgggtgatg	gtattcttct	60
taacattttc	tggctccac	ggatgtgttc	cgacatctca	gccctggaag	gatgctgaag	120
accaggttgg	gtgtgtccat	gccgtagctt	gggtgaactc	agctctttac	acagtccctc	180
gccccctttc	gggaaagccc	aaatgttcat	tctcatttga	taggaacgag	agtgaggatt	240
tgaataagca	ggaggTTaa	tgcagggcag	tgccgtgtct	tgtgtcgagc	tcaatgttgt	300
aattgtgctg	tgtaaaaggc	ctgtgtggtg	aacaagggtg	agctcactcc	agggaggaga	360
aggactgtta	gaagactttt	gtggcacctg	acagccctgt	ggggtcagct	tattctctcg	420
taccttgaac	aacttggTcc	taaggccctag	tagagatttg	aaggaaagaa	gcaacccagt	480
cctcaactct	gcttttttta	gaatgaagaa	cagactagca	aaatagcatt	gccatacatc	540
tcaaggcaga	gagatgcgac	agggaattga	agccaggtaa	ttggtcagga	aacattcttg	600
agacaaatTT	ggggaccaag	actcaaggat	tgggaaggac	aaggaaatag	gatctagggtg	660
gtctaccgct	taggcctgtt	ggttctccct	tctccatgat	agttagtggg	gaaatcccac	720
gtaaggaaa	cacgggtagt	aagaaacttg	ggaacaaata	acacctagaa	actgaggcag	780
caagatgcac	cttagtctag	gaagccttct	tgaagagggg	gagttctctg	taagaatttg	840
aaagaaaaga	aatatggctt	gcttagcaag	aatataagaa	aggctttgag	gaagaaaaga	900
tagccagtga	gtgccaagca	tctggttggg	cttgaggggtg	agcacaacaa	ggaagcaacc	960
cggccagccc	ctctgtgttt	ctgccacagt	caaacagtgc	tcaaggaaata	tgaatacggc	1020
tgctctgatt	gtgaaaaga	agagggggccc	gaggcaagg	aagctggcag	gcagctctctg	1080
ctgatcctcc	agatgctagt	tgataaaggc	ccaatttcaa	atgaaggttt	tgaagcaaaa	1140
aggacagtga	gaacccggag	gcagggaatg	aatcacagga	cttgggagcg	ggtgtgggggt	1200
gaacctgaaa	ttgagacagg	attaaaaacg	acctgtctga	gatgggacag	gggctggctt	1260
gtttcacgga	cttcaatgct	tctggcagca	atggggaaat	tgggcaggct	ggctatcata	1320
ggaggctggg	cacagaccct	gagcccaggg	gatggtacat	tgaatagcca	gtggcccccg	1380
gtgaaagttc	tgcagccaaa	aacaactggg	ggatgaggaa	aaaaggaaaa	attcaattct	1440
agtctctccc	attaagcccc	cttcccaatt	tgaagactgg	cccaagaggc	cttcgggaat	1500
acccctctct	tcttccaccc	ttctcatcac	ttccctgtcc	cttctctgtc	ctttccccca	1560
actctccccc	tcaagcccag	tctcgttgtc	accaaggctt	ctaggtgatt	agagaatccc	1620
acctcatctc	cacctggaac	cctccctcca	cttctgcact	cctagggata	aaccgttgca	1680
caccctgccc	ccacctggaa	gggcctacag	ggtctccagt	gaaaaacctg	tgaactgttg	1740
aacctctctg	ttggtggcat	attatttttg	tttttggtga	ctttttcttg	gaataagtca	1800
acaaatatta	accaagtgc	taccacatgc	caagcgtgc	tctaggata	cagtgggtgag	1860
caaaagtggg	ttgagttttt	caatagaaaa	tccatgtttg	ggtaatttaa	gcttaaaaaa	1920
tcatgc aaac	aggctggatg	cattggctca	cacctgtggt	cctagtactt	tgggaggccg	1980
aggcagacag	atcacttgag	gtcaggagtt	caagactagc	ctggccaaca	tggcgaaaca	2040
ctgtctctac	taaaaaata	caaaaattag	cggacgtgg	tggcggggcg	ctgtaatccc	2100

```

agctacccgg gaggtctagg gatgagaatc gcttgaaccc aggagtcgga ggttgccagtg 2160
agccgagatc ccgccactgc actccagtat gggcagcaga atgagactcc atctcaaaaa 2220
aaaaaaaaaa aaa 2233

```

```

<210> 57
<211> 1963
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> SITE
<222> (1540)
<223> n equals a,t,g, or c

```

```

<220>
<221> SITE
<222> (1935)
<223> n equals a,t,g, or c

```

```

<400> 57
ggcacgagct ttgaagagag agttcaagag ggcgtcatct acccttccat gtgctggatc 60
cgggactccc tggtcagcta catcaccaac ctgggcctct tcagcctggt gttctgttct 120
aacatggcca tgctagccac catggtggtg cagatcctgc ggctgcgccc ccacacccaa 180
aagtgggtcac atgtgctgac actgctgggc ctcagcctgg tccttgccct gccctggggc 240
ttgatcttct tctcctttgc ttctggcacc ttccagcttg tcgtcctcta ccttttcagc 300
atcatcactt ccttccaagg ctctctcacc ttcatctggt actggtccat gcggtgcag 360
gcccggggtg gccctcccc tctgaagagc aactcagaca gcgccaggct ccccatcagc 420
tcgggcagca cctcgtccag ccgcacttag gcctccagcc cactgcccc tgatgatgaag 480
cagagatgcg gcctcgtcgm aactgcctg tggccccga gccmggccca gccccaggcc 540
agtcagccgc agacttttga aagcccaacg accatggaga gatgggcccgt tgccatgggtg 600
gacggaytcc cgggctgggc ttttgaattg gscttgggga ctactcgggt ctcactcagc 660
tcccacggga ctcagaagtg cgcgcctatg ctgcttaggg tactgtcccc acatctgtcc 720
caacccagct ggaggccttg tctctcctta yaacccctg gccagccct cattgctggg 780
ggccaggcct tggatcttga gggctctggca catccttaat cctgtgcccc tgcctgggac 840
agaaatgttg ctccagttgc tctgtctctc gtggtcacc tgagggcact ctgcatcctc 900
tgtcatttta acctcagggt gcacccaggg cgaatggggc ccagggcaga ccttcagggc 960
cagagccctg gcggaggaga ggcctttgc caggagcaca gcagcagctc gcctacctct 1020
gagcccaagg cccctccctc cctcagcccc ccagtcctcc ctccatcttc cctggggttc 1080
tctcctctc ccagggcctc cttgtcctt cgttcacagc tgggggtccc cgattccaat 1140
gtgtttttt ggggagtggt ttccaggagc tgcttggtgt ctgctgtaa tgtttgtcta 1200
ctgcacaagc ctcggcctgc ccctgagcca ggctcggtac cgatgcgtgg gctgggctag 1260
gtcctctgtt ccatctgggc ctttgtatga gctgcattgc ccttgctcac cctgaccaag 1320
cacacgcctc agaggggccc tcagcctctc ctgaagccct ctgttgcaa gaactgtgga 1380
ccatgccagt cccgtctggt ttccatccca cactccaag gactgagact gacctcctct 1440
ggtgacactg gctagrgcc tgacactctc ctaagaggtt ctctccaag ccccaaatag 1500
ctccaggcgc cctcgccgc ccatcatggt taattctgtt ccaacaaaca cacacgggta 1560
gattgctggc ctgttgtagg tggtagggac acagatgacc gacctggtca ctctcctgc 1620
caacattcag tctggtatgt gaggcgtgcg tgaagcaaga actcctggag ctacagggac 1680
agggagccat cattcctgcc tgggaatcct ggaagar*tc ctgcaggagt cagcgttcaa 1740
tcttgacctt gaagatggga aggatgttct tttacgtac caattctttt gtcttttgat 1800
attaaaaaga agtacatggt cattgtagag aatttggaaa ctgtagaaga gaatcaagaa 1860
gaaaaataaa aatcagctgt tgtaatcacc tagcaaaaaa aaaaaaaaaa aaaaccggca 1920
cgaggggggg ccgntaccc aattcggcct ttggaaatga gat 1963

```

```

<210> 58
<211> 1267
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> SITE
<222> (1248)

```

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1255)

<223> n equals a,t,g, or c

<400> 58

gctgcagcag	actatgcaag	ccatgctgca	ctttgggggc	cggctggccc	agagccttcg	60
ggggacttcc	aaggaagctg	cttcagaccc	ctctgactct	ccaaaccttc	ccacaccagg	120
gagctggttg	gagcagttga	cccaggcctc	ccgggtctat	gcctctgggg	gcactgaggg	180
ctttcctctt	tcccgatggg	caccggggcg	tcattgggact	gcagctgaag	aaggtgcaca	240
ggagagaccc	ctgcccacag	atgagatggc	accaggcagg	ggcctctggt	tgggaagact	300
atttggagtg	cctggggggc	ccgcagaaaa	tgagaatgga	gccctaaagt	ccaggagacc	360
atctagctgg	ctgcccccca	cagttagtgt	gttggctctt	gtgaagcggg	gggcacctcc	420
cgagatgcct	tctcctcagg	agcttgaggc	ctcagcacc	aggatggtgc	aaacccatag	480
ggcagtgccg	gctctctgtg	atcacactgc	tgcaagacct	gaccagttga	gcttccggcg	540
tggggaagtg	ctgcgtgtca	tcaccacagt	ggatgaggac	tggctccgct	gtgggcggga	600
tggcatggag	ggctctggtg	ctgtggggta	tacctccctt	gttctgtagc	cctgggaccc	660
tttctcgctg	atgtgtcttc	ttcctgtcac	ctgggaatgg	aatggccagt	gaacaccatc	720
ccagaagcat	tttccctctg	caaaatgacg	tttcttccca	cgtctgttcc	tgctaataat	780
taaaataaac	tttctctctt	ccctcctata	cccacctgta	aggtgaaatc	tgctcttctt	840
ccaaatatat	aaaaaaggaa	ttgcctccca	ggtaatccct	ttcctttttc	ccgtctatat	900
aagggaatgt	cttctctcct	atctatctgc	aaaatggaaa	tctagacctc	cttcttcata	960
cataagtggg	ctgtgccagt	acaatacatg	cctcagcccc	caagcctaga	aggacctcta	1020
gtctctcttc	tgtgtggaat	cttccccact	ccatccctcc	caagttgcct	gtattgataa	1080
tgtactcact	catgctgtac	taggtgtctg	agcctggaca	cccttggtgg	gtgggcctgt	1140
ggtgatgggt	tgcacccctc	ctcctttgtc	ccaataaagt	atgggagttg	aaaaaaaaaa	1200
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaackc	gcggccgnaa	gcttntttcc	1260
ctttagt						1267

<210> 59

<211> 1295

<212> DNA

<213> Homo sapiens

<400> 59

ggcacgagct	tgtcccaggm	ctggagcagc	tgtggagaaa	ctgggaggga	agcccgtcca	60
gctgatttcc	aagccacat	gctgctcaca	ggtcaaggcc	gagggactga	tttttgccgg	120
tctgactgga	ctcaagttac	ttcccagttc	cttgacagaga	gctgtctttg	tgagacagtg	180
tcttgggttc	tggaaatgat	ggagccgtgc	tttgcaaatg	aggagtgaat	gcgtgtctcat	240
ctggcagctg	tgggtgtgct	tgtgtggcatc	aggcctgagc	ggtgaccgtg	ctcctctgat	300
tgtcctcact	gcgtgtgaca	aggcctgggc	cactgtgtga	gtcgtcttgc	gctccatgaa	360
gcctggtgtc	tgtgcagatg	tgtgggtggc	gttaagggtg	ggggacattt	gtctttcaca	420
ctggagaatg	ggagtctgga	gctggtgcta	ctggtgagga	agaggcccg	cctgctgcca	480
ggttcgcccc	caccttcccc	ctggttgttg	ggaaaaccaa	ccttggaatg	gccaaggcag	540
gagatagcac	ctccccgggtg	aagatccagg	agctctcatg	agctccacgt	ggaaagatca	600
aggatctgga	gtctggagcc	cttcaggcag	caactcagtg	accatgaacc	tcagctctgt	660
ccaccgggca	cagcattgct	gggagctgga	cccgggaggc	tgccggctcc	agagttagga	720
gggtccagac	catgcagaca	atatgccctt	tttctccaaa	caccatttca	agcaaaccgg	780
cagggtctcct	caacggctgt	cagcagcttc	tcgtagagct	tctcatagga	ctcatatggt	840
ggaatgtcga	tccggttaaa	gctgaaaaag	gacaaaagag	agtcaccgtg	tgggcagttc	900
agccctagga	ccaacctcaa	ggccaaggac	aggcagtgag	aaagacaggg	tctcgctagg	960
ttgcccaggc	tgtctcaaa	ctcctggcct	caagtgaacc	tcctgccttg	ccctctcaac	1020
gtgctgggag	ccactgtgcc	caatcaacac	acagtaaagg	ggaagctcat	ttccagtatt	1080
tgtgcaaaga	aaaagacatc	ctttaagaag	ctatcgtagc	aaacaaaaaa	atacaaaatt	1140
gtgacccaga	ggatgtacag	tgacttcttg	ctttctaggg	tgtgtgggca	ggtgctgtgg	1200
cttttgagtt	ctgatgatga	caaaaatatt	ttggcagaga	ctccatctca	aaaaaaaaaa	1260
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaac	tcgag			1295

<210> 60

<211> 915

<212> DNA
<213> Homo sapiens

<400> 60
 acccagcggt ccgtgttcac agacagtagt ttcaaagtgt gtaccacatg aagttgcagt 60
 cttccaacct tccagccagt gtgtatggaa ataacctgaa ttgtattaat agcagttctt 120
 caatgtgggg ctgctggggg atgcttggtt gtattccgtt gtttggtccc tgggtgcccc 180
 tcttgggcaa gcattttctt ggatgtcyct atttatgtgg caggtmaccc tgctggattg 240
 ccttcacatg tgtgcgact ccctgtggac caactacagc gcctacagct actttgaagt 300
 ggtcaccatt kgygacttga taatgawcct cgccttttac ctggtccacc tcttccgctt 360
 ctaccgctg ctcacctgta tcagctggcc cctgtcggta agagagtggg ctggccctgt 420
 cctccgcatg cacaagtcag gatgttagct agagtactga gacctgacag agtttttccc 480
 gtcctgccat ctcacctctt taaccattct ttgctgcctc tgcctgaat ttcctattgt 540
 ttggtggaca tctctgcttg atgtcctgct ggtttttaaa actcacttc cagctacaag 600
 aaggctgtgg ctggccgggc gcggtggctc acgctggtaa tcccagcact ttgggaggct 660
 gaggcggcgg gatcacgagg tcaggagttc gagaccacgg tgaaccgccg tctctactaa 720
 aaaatacaaa aaatcagccg ggcgtgggtg cgggtgcctg tagtcccagc tactcagaga 780
 ggctgaggca ggagaatggc gtgaaccggg gaggcggagc ttgcagtgag ccgagatcga 840
 gccactgcac tccagcctgg gtgacagagc gagactctc tcaaaaaaaaa aaaaaaaaaa 900
 aaaaaggcgg gccgc 915

<210> 61
 <211> 1445
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1047)
 <223> n equals a,t,g, or c

<400> 61
 aggaattcgg cagcagcggc acgaggactc cttctcttct gcagaagcag atgggaatat 60
 gctcttttaa actatgagat actggacaga catgaggagg aactaccgtg tcacgtatca 120
 agtagtgttg ttatttctgt gcttctccct cctaacagaa tgtaaaacct ttgaaccagg 180
 gtccagagagg tctttatttt catatcccct gtgatgtcta atttatttgg atttacagat 240
 aaatgatcgg taaacttttag aaacagcact ccagtttata gctctgtgct gtagacttac 300
 tgaacaacta cagtgaacc aattcaaaaa gggatatttt gtattatgat ttagtctcct 360
 acttccaagg ctagttttta aggctgtgaa ggaagctga aaatgacaca gtgtttctgg 420
 gatgaccaga cagacactgt atccagagat gctgtctgag cagcggggga tagtaaaccc 480
 cttagtacaa cattaatttg catggtggtt tatgagttaa tgtaatacca aatattaaca 540
 taaataaaaa tatatttaag tgataactaa gctggacata tatcttaaaa gacaactaca 600
 gcccagaaaa caatgaacat tgttgcctta cagctatttt gtcactgtga tgatacctaa 660
 ttttaattct aaaggagct gatgtttata acctagaagt tgattttgat aacatttgag 720
 aaaacttcat aaagctggca caggtaacat atttagtttt gtatatctgc tgtccaattt 780
 gagtctctaa aaattatctt agaattgaata tgaattcgc aggtataaag accaagtttt 840
 cagaaaataa aaatgtccaa gtactttgaa acatctattt ttcactcatt attcagccta 900
 ggatattagc acttgtgtcc ttgaacagag atgagaatgt ttgttatcca aagaccagga 960
 aggtcaccag ccaagggata tacagtcgtg cctcatcttc tgtgectttg tattccttta 1020
 tgctttgtag cttaacaaaa ggttttncct tgtaactgtt aagtttccat atatttgta 1080
 aatataact tcacacttca cagttgctca tgtcagaaca gactattgaa aatgtaaac 1140
 tggccaggca cgggtgctcac gcctgtaac ccagcacatt gggaggctga ggcaggcggg 1200
 tcaactgagg tcaggagttt gagaccagcc tggccaacat ggtgaaacct tgtatctgct 1260
 aaaaatgcca aaaaattagc taggcatagt ggtgcacgcc tataacccca gctacttggg 1320
 aggctgaggg aggaagaatt ctgaaccaca ggaggcggag gttgcagtga accaagatca 1380
 caccactgca ctccagccwa ggtgatagag tgacactctc tcaaaaaaaaa aaaaaaaaaa 1440
 ctcca 1445

<210> 62
 <211> 1100
 <212> DNA
 <213> Homo sapiens

<400> 62
 ggtgactgct ccctagctgg tcatgaaaat tctctcaag attattaaat cagggattat 60
 gtcttgtcca aatataagtg aaatattgtt tgtaacaatg ataagttact tggctttaca 120
 ttttagtaac taccctttca tgtttcttta actcttgaaa tattttatta ggggttgagc 180
 attcatgatg gtacctggaa gtcagcaatt tatggttttg gagatcagag taatttgaga 240
 aaactaagaa atgtatcaaa tctgaaacct gtcccgtca ttggtccaaa attgaagaga 300
 aggtggccaa tttcttattg tcgggaactc aaaggttatt ccattccttt tatgggatct 360
 gatgtgtctg ttgtaaggag gactcaacgt tacttgtatg aaaatttaga ggaatcacca 420
 gttcagtatg ctgctgtatg aactgtggga ggcacacct ctgttattaa gctgatgttt 480
 gcaggacttt tctttttgtt ctttgtgagg ttggaattg gaaggcaact tctcataaaa 540
 ttcccatggt tcttctcctt tggctatttt tcaaaacaag gcccaacaca aaaacagatt 600
 gatgtgcct cattcacgct gacattcttt ggtcaaggat acagccaagg cactggtaca 660
 gataagaaca aaccaaatat caaaatttgt actcagggtg aaggaccaga ggctggctat 720
 gtggctaccc ccatagctat ggttcaggca gccatgactc ttctaagtga tgccttctcat 780
 ctgctaagg cgggcggggt cttcacacct ggagcagctt ttcccaaac aaagttgatt 840
 gacagactca acaaacacgg tattgagttt agtgttatta gcagctctga agtctaaca 900
 ctggaagaat taactgaagt cataacgtgc gtgaattaac agcttctcta ttgatattt 960
 gaaattcttc tgtaagcctg tctgagtgtg tgtggaaacg attgtcaaat ctaaaatatc 1020
 tataatattaa aaagtaggaa attgtcctag cttaccctaa atttcaaaaa aaaaaaaa 1080
 aaaaaaaaaa gggcgccgcg 1100

<210> 63
 <211> 1499
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (52)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (66)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (84)
 <223> n equals a,t,g, or c

<400> 63
 agcttattgc aaagacaaat gtttgaagtg tttgttgaga ttctctgttg tnttctctga 60
 ggcagncaca gcataagctc tttnacccctc tacttctcag cacataagct ttcttaccat 120
 ctatcactgg agtcaggggt gaggggagga ccgcatgaca gttggttaat atacacttat 180
 tttttggcaa aaacgttttc tctgggacca gaatgatctt gatactgaaa aaatttctag 240
 tgcctagatcc tctttctaag tgtgaaagga cttatcttga atgctccaga atgatcccaa 300
 gtgttgagct gagagggacc tggcagcaga atctgattat tgaaaagtgg caattgttga 360
 tttattgaag acagaataat aactcagcag aactgttatg ttgagctgaa cccgacctcc 420
 ttacagccgaa tcatgcaaga atgcctgctg catggctgtt gctgctactt attaaggctt 480
 ggtgttcttg gcacagtgc aatgcatttct acatgggttg tcttcacagc aaatgaacaa 540
 cacaggctta aggaacaaga caactctcaa agtctctcag tgagtagagc ttagctgttg 600
 gtagtcaaca tgccacgcga ttccggragt gagcctgtct ccagagggtta gagatgttca 660
 gtttctcttt aaggttctta cgtagatttt ttccatgact ttatctacat cctccttaaa 720
 tttacgtttt tagtccctac tggctcttga tatcaccagt tttgttggtta ttagtaattt 780
 ctaactgccc taaatttgct tgttttaaga ttcaagggtg gatacctcag tctgttatct 840
 ggaatatggt ttacaaatcc attttttctc ttcaaggctt tgaaaacatt gacattgtct 900
 cctcctaaca tttttatttg tcttgccagc tcttaattta ttttaattat cgttaggaag 960
 acgacttttc tgtcttttga tgatttttagc tgcccttctc tagaccttgc tgattccatt 1020
 atctttacca agaattgaaa gtgaaagtgg catttgtcat agaattgccat ggtcttatct 1080
 caaagtatct taggatggaa caatacaagg cataatatgg ggtcagttag gttgtttaca 1140
 cgagtgaatg accaacaaca ctactgtctg ttcaaaccga gtctgaaggg tgaatcagac 1200

cgaccattgg	ccgtgagggt	ctggactgct	cagtattatc	tcaaggatat	caagggttat	1260
tggaaactgt	gtgatcaaag	gggctccatg	actttatgca	gggattcagt	agggagccaa	1320
gaagggttag	aatagttcag	agaccagagt	ctaagaccaa	tcaagaagaa	tggatcaatt	1380
agagatatga	attctgggtg	ttatatTTTT	gtggagctgg	ttgtgagata	aaagggtcaag	1440
cctaccagac	tgaaaagtgt	atgtgaaagc	tctttaaaaa	aaaaaaaaaa	aaactcgag	1499

<210> 64
 <211> 655
 <212> DNA
 <213> Homo sapiens

<400> 64						
ggcacgagggc	aggaaccgct	aaacgagaca	gacactggcg	actcagagcc	ccggatgtgt	60
gggttccttt	ctctgcagat	catggggccc	ttgattgtgc	ttgtgggatt	gtgtttcttc	120
gtggttgccc	atgttaaagaa	gagaaacacg	ctgaatgctg	gccaggatgc	ctctgagaga	180
gaagagggac	agatccagat	tatggagcct	gtccaggtca	ctgtagggtga	ctcggtataa	240
atatttccac	cccctccacc	accttacttt	cctgaatcct	cagcttctgc	ggtcgctgag	300
agtcctggaa	ctaacagtct	gcttccgaat	gaaaaccccc	cttcatatta	cagtattttc	360
aactatggga	ccccaacttc	agagggtgca	gcctctgaaa	gagactgtga	atctatatat	420
accatttctg	ggacgaattc	atcttctgag	gcctcacaca	ctccacatct	tccatctgaa	480
ttgcctccta	gatatgaaga	aaaagaaaat	gctgcagcta	cattcttgcc	tctatcttct	540
gagccttccc	caccgtaaac	tatggactct	agttcagttt	tatatgcaat	ggatcactac	600
tccatcaatt	tcttcaaaac	aaaaaacac	agcaaaaaaa	aaaaaaaaaa	aaaaa	655

<210> 65
 <211> 1446
 <212> DNA
 <213> Homo sapiens

<400> 65						
ggcacgagcg	gaagtgcac	tcgaacttgg	tcggggcgcg	gatccccgaga	gggaaagtca	60
taacaaccgc	acgagggagt	tcgactggcg	aactggaagg	ccacgcctcc	tcccgcctgc	120
cccctcagcc	ctgtggctgg	ggcagagctc	agactgtctt	ctgaagattg	atgtctatgt	180
ccttgagctc	tttaattttg	ttgccaat	ggataaacat	ggcacaaatc	cagcagggag	240
gtccagatga	aaaagaaaag	actaccgcac	tgaaagattt	attatctagg	atagatttgg	300
atgaactaat	gaaaaaagat	gaaccgcctc	ttgatttctt	gataccctgg	aaggtttgaa	360
tatgctttta	atgaaaagg	acagttaaga	cacataaaaa	ctggggaacc	atttgttttt	420
aactaccggg	aagatttaca	cagatggaa	cagaaaagat	acgaggctct	aggagagatc	480
atcacgaata	tgtatatgag	ctcctggaaa	aggattgtaa	tttgaaaaaa	gtatctattc	540
cagtagatgc	cactgagagt	gaaccaaaga	gttttatctt	tatgagttag	gatgctttga	600
caaatccaca	gaagctgatg	gttttaattc	atggtagtgg	tggtgtcagg	gcagggcagt	660
gggctagaag	acttattata	aatgaagatc	tggacagtgg	cacacagata	ccgtttatta	720
aaagagctgt	ggctgaagga	tatggagtaa	tagtactaaa	tcccaatgaa	aactatattg	780
aagtagaaaa	gccgaagata	cacgtacagt	catcatctga	tagttcagat	gaaccagcag	840
aaaaacggga	aagaaaagat	aaagtttcta	aagaaaacaa	gaaccgacgt	gatttctatg	900
agaactatcg	taacccccaa	agagaaaaag	aaaggatgca	attgtatatc	agagaaaatg	960
gttctcctga	agaacatgca	atctatgttt	gggatcattt	catagctcag	gctgctgctg	1020
agaatgtgtt	tttcgttgct	cacagctatg	gaggacttgc	tttttttgaa	ctgcaactca	1080
tgatcaaaac	agctaattca	gatgctggga	agtgccttcg	cttagctatg	tggaagaacc	1140
attgactgta	tacaaccaac	aagtgtatgg	tgcaacagga	gatccattga	aaaccgttta	1200
taggactgaa	cgacaacccc	aatgcaagt	gaccatgagc	aactacaaat	aggtatacat	1260
atgcatttga	gctgaacaga	ctttctgaca	tataatttag	tcaaaattgc	tgtatttctt	1320
ccccttaaat	ttatacataa	tcagcttctt	gtatggaccc	aaattggaga	aatgtaattc	1380
agtagttggt	gagaaataaa	ggattgtgac	ctctgtgtaa	ttatcaggaa	aaaaaaaaaa	1440
aaaaaa						1446

<210> 66
 <211> 670
 <212> DNA
 <213> Homo sapiens

<400> 66
 ggcacgagag ggcctaaggg gaacaccccc ttccccaggt cttttatttg ttttaagttat 60
 ttttgcacaa atgactcttt tatatttaat tgcatttcat tgccctccct ctttaagcca 120
 acaggctcag tttacaaacc tgtgagctac tgttggctgc tgccctccct cccagtgaag 180
 ggtacaaagc aataagcatc atgcatcctc cccttaccct tccaacaccc ctctgcctct 240
 ggctcagggt gctcaaagca cagatcctct cttaccccg cccaggttt gaaacacata 300
 gcctcatttc aaggtgtagc caggttcccc cgactttcct ctgggatata aaaaaggggg 360
 taagggggc aagagagccc tctgggctc tcctcccata cactactac tgcccttct 420
 cccccatca aaacgctcag agacgtttgt atgatgcgc tgaggattat gcaacgtggt 480
 ccaaccggag cggccagcat gaccagctgt ccaggggctg cctcctgcct tttctttgt 540
 aaagacaaga cccttgggag ttttaattct gttttgtact tgccctgtgg ggcctccact 600
 gcttttctat gggagacact ctttaattta cagatgagaa tattttgaaa aaaaaaaaaa 660
 aaaaaaaaaa 670

<210> 67
 <211> 1692
 <212> DNA
 <213> Homo sapiens

<400> 67
 tgcagtccca gctactgggg aggtggaggc tgcagtgagc cgagatcaca ccactgcact 60
 acagcctggg cgacagagag agactctccc aaacaaacaa acaaaaccc aaaaataaag 120
 aagtcactct gaaagaagtt tcaacatttg ctttttcatt ctgagattac agttttctat 180
 aaacatctaa gagtgaagag tctgacgttt tttggtcaca gctgagccac tgcgtgaccc 240
 ccgccccgcc ccacactcac tttgctctag gcaaaagctgt actctgaaag ctggccccaa 300
 tggggagggt aggactgtgc ctgctcagaa gtctgtgggt gcctcagaga agggcaacaa 360
 ccctaggctg gaccctagcc ttgagagtag ttctactgac cagagccsc agatyycttc 420
 cgggtggcgc agatactgcc agaagagcct gcggtgcaca caccagaatc cgggtacttg 480
 gatgagaagg acacattact gatcaccttc ctccaggcaa ccctgtcagt taaggactac 540
 agtccccgcc ccattatgta gatagggaaa cagaggcaaa gaagttagga aactcgccca 600
 gaactctcag ctcatgaata aaaaagcaga actaaaaccc agtgctctcc ctggctgggc 660
 aaacgtgtgg aagttgatgt gcctgggttac tgtttgtgct tcgcttatca taaccagtga 720
 cagcgtggtt agcactgttc gcctcaaggg cagctgtgag gattacttgg gattgtcctg 780
 tggaaacact tcacatgcat attaaactagg agaaaagcca ctggagaatg agctttatga 840
 gctctatcaa tcaccacagc tagtctgacc taggggtaag caaaatggaa gacaggaaaa 900
 agggaaataca tttgctyagg acagcgtgag ggccacgtga gctgcttgat tggtagcgat 960
 ttgtacaggg gctttatgga tcaactaggt ttaatttgca aggcctgaaa ctgtccttag 1020
 cattctctga aacccacagt gccagtgcct ctccacgctc cgccagcag aaagctcctc 1080
 atgactggat cctcttgaga acttcagagg ggtcagtgga cggtgactga gactgcctca 1140
 gtgatcacgc tcggtgctat gagctgaaat ctgggccaag ggcacagtaa gtccaggcag 1200
 ctagtatgtt taaaataact acttttcggg agctaagcca tgaggacgta aaggcattaa 1260
 gaatgatata atggactttg gggactcagg ggaaggggtt ggggtgaggg ataaaagggt 1320
 ccagtgtaca ctgcttgggt gatgggtgcc ccaaaatcct ggaaatcacc gctaaagaac 1380
 ctacagtaac caaacaccac ctgaacccca aaaacctact gaaactttta aaaattaaaa 1440
 atacatacat aaaatagcta cttttactgc tgtcaacagc atgttctcga aaaatgttgg 1500
 aattcaaaact ttctggaggg cagctgtgca agaaacttat tcacgtcagg agttttctaa 1560
 aatttgtttt taatgcttat tggtaactct gcattagaag taactacaaa tgtcttatta 1620
 aagtttccac ttttaaatgca aaaaaaaaaa aaaaaaatga ccctcgaggg ggggcccggg 1680
 acccaattcg cc 1692

<210> 68
 <211> 655
 <212> DNA
 <213> Homo sapiens

<400> 68
 gatgtagagc agactgagct catccatcat gatttcttcg tgatattact gccaaagcaga 60
 ttataagggt aagtcfaatgt gacaaaagga aattcggcta aaagcttctt gaagcctttt 120
 gatgctaagc agtccctctt ttgatattta ataccatgg acataaaact ctgccttaga 180
 ggtcgccatg gagttttgtt ttgttttgtt ttgttttgtt tttgccatct gttaacagtc 240
 ctgagtaccc atagagcctt ttactattta tcagcatyct agagtcgtca gtatggattg 300
 tcaaaacttg cattkgtctc ttttttgttc agtgttgtgt gcattccatc tttctttctt 360

ttttaaaca	ccctgcttat	gtaacatcca	cattttctga	cttacctttc	aaacctgcc	420
gaaagcagaa	gtgatattta	awacacttgg	taagtgtttat	atatwgattc	taatgataat	480
gttttrgtcta	agatggacct	gacaaggcca	ggcatrgtgg	ttcaacagca	ctttgagagg	540
ctgaggcagg	atgattgcct	gagcctggga	gttcaagggt	acagtgaact	gtgatcacat	600
cctgccttct	agcctgggtg	acagagcaag	accctgtctc	aaaaaaaaa	aaaaa	655

<210> 69
<211> 1618
<212> DNA
<213> Homo sapiens

<400> 69	
taacgcgcct	gcaggctcgac
aaaaaaaaaag	aataactgcct
taacatgcag	tcacatcaaa
ataaataaaa	tggtttctgct
gttgcaaaaa	gtttcagagg
gcttttatct	ttgaaattgg
ggaatagaaa	catgaaaccc
aaatgagaac	acttgaatta
tcaagaccta	ttgtgtcgct
tagcaaatcc	cctaaatacc
atcactctcg	tcaaagtcaa
tctgtgcctc	ggcttttact
tgtattggag	acaaagagag
ctatctgcag	gttccacatc
agtatataaa	aacaaaataa
gcattttacat	tgtatttagat
tgaacaacat	gaggggttagg
aacttttgag	ttcccaaaaa
aatacagtca	attaacacat
tgaagtaagc	tagagaaaaat
ctttttttaga	gagagaggttc
agcaatcctt	ctgtctctgc
cagctcctat	atattattatt
aggtggagta	ggctgaggag
ggcagaggca	gaagaaagta
tcagctgttaa	acaggagggc
gggacttggg	catccatgaa
	ttttggcatt
	tagaggttcc
	tggaaaccaat
	ccctcgag
	60
	120
	180
	240
	300
	360
	420
	480
	540
	600
	660
	720
	780
	840
	900
	960
	1020
	1080
	1140
	1200
	1260
	1320
	1380
	1440
	1500
	1560
	1618

<210> 70
<211> 1802
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1790)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1792)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1801)
<223> n equals a,t,g, or c

<400> 70	
gaattcggca	cgagtccttc
tcacttttga	aatgcttatt
attttaatga	caataatgca
	60


```

gagagagaga gtatttttga atagacttaa gttttccttc aactaatgtc tccttgagg 120
acagaaatac aactaaaccc tctgtcaacg tgggtatgta tttttttact ttctattttt 180
caattagttc ttttatgttt tttcttctag tcattgttaa agctaccaat ggaccaagat 240
atgttggtgg ttgtcgtcga caggtaatac tttatatatt tatagtgcct gatgattgac 300
aaagcagttt catgtaagtk attgtctcya attcttgagg cwagcagggt gagcattttat 360
gcccataact cacaaggatg atttgttcag acatagctag ttattaacaa agcctgaatt 420
caamccatgg gctttgactc ctggcattcc gtactttcta ctgtattaca ttgtctcagt 480
cagatctggt aatagccact tagaaataaa agtatttttag aactggaaaa cagacatttt 540
attttaatgt catttttaaa gaggacttaa aagtgttaga tatcatcagt tacctgtgtt 600
tatatttaga cattcagaac tgttacttat ggactgtacc atggcctaag ttaattttgt 660
atgaggtcat ttagattagg gttagggcaag ttgaaataat tctaaatttt attttacagt 720
tatcaaagat gccacaaaat gacctcaagt cattcagtag tgtctgaaat caatttatgt 780
attattcttt aggaagtgtc cttagataat tcttttaaat tcattggaag agttttctct 840
gtttaattgt catttcaggt tcaggtttta aaacattcac agaacatggc tgaagggag 900
aatttaatcc aggaactata aatctcctat taggattttg cctagtatat aagcggttga 960
cattttctaa gtcaaaatat tagataccta aactgacaag ggattttcat gtccctttca 1020
gggctctgtg gatgccgaaa gttggcattt ctaagatatt tcaggttgca tgaggacaag 1080
actgtatttg aagactaaaa aacattagaa aagccgaagt atataaagt tgagtatccc 1140
ttatccaaaa tgcttgagcc agaaatgtgt tttagatttt ggcttttttt ttttcaggtt 1200
ttagaatatt tgtgktgkac tgggtgagca tyoctaatta aaaaaaatca aaagtttgaa 1260
atgctccgag gagcattttt tttagcagtc atgtcagcat tcaaaaaatt tcacattgkg 1320
gagcattttg gattttcaga ttaagaatac tcagcctgka ttctctatag atgtaaacat 1380
tgaaatagct tcatattgat ttctcctctt attttttcaa gtaacctcac ttcttagccg 1440
ttttttcctt aattgttata ttaactctag tgttttgctt atcttcttaa atttgaagct 1500
ctttgtaaaa tctgttgaca agtggtcagt aatttatatg attccgaaat tgtattggca 1560
cgcagttttt taaactatta aaaagtaact tgggtcgggc ggggtgggct atgcctgtaa 1620
tcccagcact ttgggaggct gaggtgggca gatcacgagg tcaggagatc aagaccagcc 1680
tgaccaacat ggtgaaaccc cgtctttact aaaaaaaaaa aaaaaaaaaa 1740
aaaaaaaaaa aaactcgagg gggggcccggt acccaattcg ccctatagtn antatagtga 1800
nt 1802

```

<210> 71
 <211> 1292
 <212> DNA
 <213> Homo sapiens

```

<400> 71
ggatgataga tgatctgtaa atattttctc tcattccata ttcctctggc tccttttaag 60
attttttttt atgtctggtt tcaaggaatt tgattttggt gtgccctggt ggagtataag 120
ctttcttttg agtttttttg tcttggtgtt aagcttttgg agtttggtgg ttataacttt 180
tcacagattt tggaaacatc ttggctatta ttctccaaa tagtcacaca tgcctcctcg 240
gattccagtt acatatatat tattaggttc ttgaagtgtt cccatacctt actgatgctc 300
tgctcttttt ctttgggtctt atatttgggt ttcatittgga tagtttttat ttctgtgtct 360
ttacattcac tegtctttcc ttctgctgtg tcttgactgc tgctagtccc atccaatgta 420
tttcatttat atatctataa ttgtggttt gatagaaatg cagtgatgta gcaggatca 480
ataaatactg ccttaatttg ttgcgaaaat ataacagatt cctgttctgt atgttagcta 540
aaaaggtatg caaacacccc tgtatgtcat attaacattt atgtcccttt gtttccatgt 600
caacttttag ttctctgccc aaaacctaca tatgtttttt ttatatgatt attctacatt 660
ttctgtgagc agtggacatc tgcattagta gttctatgat atttgtttta taagttgcca 720
gaatgggttg tctgtttggc agactgcaga caaatattta tctatgatcc gttgcatgat 780
atgaccatga ttttgcctaca aaaaacttga aatagatttt aatattttct ttactattat 840
cagagagaga gctggattac ctgcaaaagt gtacttttgc ttattgtgtg cattgataac 900
tcagtgcagc ctgggcgtgg tcactgggat tacctccatg tgatcacttt ttgttccacta 960
atgttaattt aaaaaatttt aggttgggag caggtgggct acacctgtaa tcccagcact 1020
ttgggaggcc gaggcagggt gatcatgagg tcaggagatc aagaccagcc tggccaacat 1080
ggtgaaaccc agtctctact gaaaatacaa aaattagcct ggcattggtg taagcgcctg 1140
ttatgccagc tacttgggag gatgaggcag gagaatcgct tgaacctggg aggtggaggt 1200
tgacgtgagc caagattgca ccattgcact ccagcctggg caacaagagc aaaactctgt 1260
ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 1292

```

<210> 72
 <211> 1054

<212> DNA

<213> Homo sapiens

<400> 72

```

ggcaccgagtg tgtgtgtgtg tgtgtgtgtg tgtctgtctg tgtgtctgtg tatgtgtatt      60
tctggctgtc tgttcattg ctctatatgt ctgtttttta tgcctgtacc atactgtttt      120
gattactggt tagtaatgta ttttgaaac aagacatgtg ggtacctcct gctttgttct      180
ccttgctcaag attattccag gtcttttgtt gcttcttcat agacgaatta actgctgatt      240
tatgaacttg aatattctga tttctttgac agttagtctt cattgtaaat tgataaatta      300
tcactctggt ttatatac acgttttttagc tatggctaatt aacagtcttt cctcacaatt      360
catatttagc atgttggcaa aatcatattt tggaacctgc aagacatagt ctctggtcta      420
tagtaaatca agctgctagg ttgtagtctg acaacttgtg taatatttta gctctggatg      480
atattaattt ttaagattat taaattttat ttttcagtgt tttacattga cagcaaaatt      540
gagtgggaag tacatactaa tttttctgta tcttagaatt tctttgggat cattttaact      600
attttaatgt tttaaatttt attgtgaatc tttttaagga aggtctgagct gttgctacaa      660
ctgtaaaata aatattctta aagcaggcag tgatgatcaa aatcttgcca ttgaccatt      720
aagctgctag aatatgagag tgataattta ggaatgagtt gattaaagaa aataacaaag      780
tagtttacta aggaattaat aatagcaaat aaaaggttta acaacaaca ataaatattc      840
tgttgatatt gcaccttaac tttccatcat catcttggga gctgactttt ttgctgattt      900
cattccgata agataagttc atttgaccac gtgattatta ttaatacat ctactgataa      960
ctctataata gaaagtggca gatttttagat aaagggtttg tgatttttaa ggttgatatt     1020
aacaggtagt atcataaaaa aaaaaaaaaa aaaa

```

r054

<210> 73

<211> 733

<212> DNA

<213> Homo sapiens

<400> 73

```

tgtgaccgat atctgcaraa ttcggcttat cgygaacctg gctttggygg acctgggact      60
ggcactcact ctcccctttt gggcagccga gtcggcactg gacttttact ggcccttcgg      120
aggtgccctc tgcaagatgg ttctgacggc cactgtcttc aacgtctatg ccagcatctt      180
cctcatcaca gcgctgagcg ttgctcgcta ctgggtggtg gccatggctg cggggccagg      240
caccacaccc tcaactcttct gggcccgaat agccacctg gcagtgtggg cggcggctgc      300
cctggtgacg gtgcccacag ctgtcttcgg ggtggarggt gargtgtgtg gtgtgacctt      360
ttgcctgctg cgtttcccc aaggtgacct gctggggggc taccagctgc agagggtggt      420
gctggctttc atggtgacct tgggcgtcat caccaccagc tacctgctgc tgcggcctt      480
cctgcagcgg cggcaacggc ggcggcagga cagcagggtc gtggcccgtc ctgtccgcat      540
cctggtggct tcttcttcc tctgctggtt tcccaaccat gtggtcactc tctgggggtg      600
cctggtgaag tttgacctgg tgccctggaa cagtactttc tatactatcc agacgtatgt      660
cttccctgtc actacttgct tggcacacag caatagctgc ctsaaccaw tagcytaygt      720
cttaagcmga att

```

733

<210> 74

<211> 785

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (716)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (731)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (772)

<223> n equals a,t,g, or c

<400> 74

ctgcaggaat	tcggcagcag	gttttatcat	ccaggatatg	gtcactctca	gtggcatatt	60
ccatgtgcat	ctgataagga	tgtatgttct	gctcttcctg	ggtaaagtgt	tataaattca	120
aattgttgat	aatgttcagg	tcacttatat	ccttaatggt	tttctccctg	attcttttat	180
taactactga	gagaagaata	ttggcatgtc	cacctataat	tttgaattcg	tctatttttc	240
tttcagatct	gtctgttttg	ccttaaacad	tccttatctt	tcagaataat	taaaagtaaa	300
aaaacattgt	tacttggttt	ttccatttct	gatgttctcc	atgttggtgc	atagatccaa	360
gtttctgagc	ttttaccctg	tgaatcatag	tcatttttaa	tttcttggtc	tatctgagag	420
tttagttctg	attactgctt	tgtcttttca	gattgtgttt	tattgtgtat	tttccatttc	480
cttgtaattt	tttatgttaa	aaaaattgtg	tatgtgcmaa	gctgaacata	ggacagaaga	540
cactgaagta	aatgttttca	tgcttggaag	tgagcaggcc	tttctctctc	ctctctttag	600
tcgtgggytt	gtgcttggtt	agttgagttg	ggtttgaagt	ttgktcacct	ttggctttgg	660
gtctcctaac	ctgaactttct	gtgtttcctg	tgcaactgtc	ccaagataga	aactgnttct	720
gggctatctt	ncagttggaa	ttccttactt	gattcttata	agcatgggtt	angaaggga	780
acatg						785

<210> 75

<211> 2341

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (161)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (163)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (170)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1229)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2243)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2309)

<223> n equals a,t,g, or c

<400> 75

gcccagttcc	tcttgaaaag	gcagagaatt	tagacagaaa	ttcaccaact	gctttcttac	60
agaaagtaaa	ccaattttct	cttccagaaa	aatggagtaa	atgtattttg	ccctattcct	120
tctactaaga	aaaactataa	accctgaaca	ttatatataa	nanatatgan	aactcagacc	180
tgagagagacc	aaggcagatg	tggtagggac	ttmataaatt	gtatagtgtat	gaatcctcta	240
agttttcttt	tctgctttat	aatttgcaga	cttttagctg	aaaatgccat	caacatagaa	300
atactaacag	gcacatatga	gaatttccca	acaaaagcct	attatttttag	gcaaagggtca	360
aggaaatagt	ctaccaaggc	agaaaacatt	tcgacaataa	ccactctact	gtagtcaagt	420
accacagaaa	acactattac	ctcaagtga	gagcttagat	cttttagayct	tcataccagc	480
caggctgtga	caagggtgtc	caacctctct	ccagaatagt	atctcagaat	agcagaagtt	540
ggaactttca	tccccaactt	gtggtaataa	gcccctcact	ctccttccac	accttgatat	600

```

gactggagag caaatgggga gctggatcta ctctaaaagc agcaatgaag aagcaccctc 660
ctttccatac caggtgggtg tgtggaggc catgtgggaa acagtaacaa gtcacttctt 720
cctccgagac aggctatcag tggaggccca gtggtgaccc agaattccacc ctccagccag 780
cagtaatgag gaacctccgc tgcctagggtg tcaacagaga ttgagaggaa acctttattt 840
ctatcatcac ctggcagtaa tgcagtgtcc ctccctcact cctttgcctt gctggagtag 900
tgtctgagga agctagctaa gacagaaaag gtaaataagt tctagagtct cataatgcct 960
aaaaatgtct ggttcattta gaaatcattt ggtatacaaa gaaccaggaa aaatctcaac 1020
ttgaatgtaa aaggttaatta gaagattcca gaacaaaaat gacaaagatg ttggaattat 1080
tcagaaaata ttttaaagca gtcatacata aaatgcttcc agtatattgs ttacaacata 1140
tatgaamcaa atttaaaaaat tatctyagcc aaaaaattaa aatatwtgaa agaactgaat 1200
ggacatttta gaactgaaac ttacaatanc cacataaaaa attcatgaag gtaagcagga 1260
aaaaactata aacacagcct cagggacctg tagtattata actgaaggcc taatttttgt 1320
gttatcagag tcccagaagg agagaagaaa tgggcaactt tgagaaaggc ctcaaagact 1380
gaaaacttcc ttaatttggc aataggcaaa aaccacrga ttcctwaatt cargcaamcc 1440
caaaatctct tagcactgta tcagaatacc atagaatggg tggtttatwa aaacaaaaat 1500
gtgttgctca caatactgga ggctggaaga ccgtgatcag aatgccagca cagatgagtt 1560
ctgtcgaaga catttttttg ctatagatgg acatcatctc attgtatcct cacatgttgg 1620
agaaaagaaa aagatatctc ttgtctcctt ctccctctct ctctctcttt ttttttttat 1680
aaggcctctg atctcaacr t gagggcccca mmctcatrac ktartctaac cctaattacc 1740
tcccaaaggc ctaacctcca aataacatca cattgaattt aggatgtcta catatgaatt 1800
ttgaggggac acaaaacttc agtgcataaa actaaccaag acaaacacaa agaatccaaa 1860
ctaagggtata ccatggtaaa atatctgaaa attaaaagaa agaacaaatt ttgaaagcag 1920
ctagagggaaa tagctcatct ataggagaga aaacaataca aatggaagca ggaaacatca 1980
gaaatagatg aaagccatag aaaagtggca caacactgtc tatgtgatga aataaaaata 2040
ctttcaattc tggtttttat atctgtgata tttgtctttt aggaatggaa gggctataaa 2100
gacatttgat gaaagaaagc tgagaggatt tgtcaccaga aggtctrcct tttaaaarrg 2160
ggctcaagar rrttctctat ccaggaaaaa aaaaagaaaa agtttaaaaa agaaacttta 2220
aaacaccaga tttaaagaaa acncagtgga aagggaaaaa tgagtggctt catcttcctt 2280
ttcctcttca gtttggtaga tttatttgn cagctgaagt taaaattatg ccattatcag 2340
a 2341

```

<210> 76
 <211> 1882
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (755)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1237)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1866)
 <223> n equals a,t,g, or c

```

<400> 76
gcaagttttg tgtttggccc tcaataaact agtctctctg tacccttggc aggggggtggg 60
aggagtcctg ggggagctcc ctccaaaatc ttacagggtg gtctgtttct tctttggata 120
ataatgatgt aatggctagt ctcttgagaa cttgtctgtt tccatacatt gtactaagca 180
tttatttggg ttatctcatt aaatcttcac aatcacttta tttaacagat ggagaaatta 240
aggcacatgg aacctaaagt tttaagggtc atggagccag taagtgttag agccaagtcg 300
tttggctcca gagcctgtgt tcttaactac tactttgtag tgtctttctt acatattagt 360
tgggctctgt tattgttagt tgaattctc tccccagtg caggccttca cgtgtttgac 420
catggttttc atgttctcca aacctcagtt ctctagattt gtactttggg aggtcatcat 480
tttccacaga tctacacct ttaggtcaga aaatcttgcc agtttataaa gattctcttg 540
gactaactcc caciaagcaa ggtcacaga gatcaatgta caaatgaagc agttcagtg 600
gtttgtctac cattctccat aagtacatgg grgacamctg atgattggaa ggtttgggtc 660

```

acctcatggg	agctgtgata	tctcactcac	cacacagatc	tgctctctcg	agggaccatc	720
ttgccaat	ccagagagtt	gcagggatat	taaanctttg	cacattaagc	ttctctcttc	780
caagctgsac	atgggscctg	ctaccgkttg	tgaamagtct	tctagagtga	tawaggttct	840
agctttctta	gttaagatcg	tattttctga	taccactccc	ttgtcacttt	gcctgaaatg	900
agaaactccc	aacctcaact	gcttttctag	tctcttccaa	tgaatgcctt	ccaaagggtt	960
gggtgtctcc	aggggtgtatt	agttgttact	aatttcatcc	tccaaggctg	atctgatttt	1020
caagatctgt	agagagacct	tagtatattg	ccttgccctg	accaaamca	gtcattatgg	1080
cmcaggaaaa	tctcaaatmc	cttattggaa	acccaggcaa	atatttattt	gaccttaatg	1140
aaatgaaaaa	gacattggat	gcatacattt	aaagaaaacc	caaaactttg	gaatctttac	1200
caaggagggt	atcttttgaa	aaggacagkc	tggaacnaag	aacttgataa	aatagaagta	1260
aagggtgaca	cttttttttt	ttttttttga	gatctatata	actctgtcgc	ccgggctgga	1320
gtgtagtggc	gtgatcttgg	ctcactgaaa	cctcggcctc	ctgggtacag	gtgattctca	1380
tgctcagacc	tcctgagtag	ctggcactat	gggcatgtgc	caccatgccc	agctaatttt	1440
kgtgtttttg	gtggagacag	ggttttaccg	tggtggctag	ctggctctga	cctcctggcc	1500
tcaagtgtac	caccgcagct	ggcctcccaa	agtgaagatc	ggcattacta	gcccctgttc	1560
gcacatgaga	cagggcactg	gatgggtgtc	acctaataat	tttcaacca	ggggcccttg	1620
gcccgaagct	atcactggta	taaaggccct	ctggcagcta	atgtgagggg	gagtgtggct	1680
gttgtttcca	tgagagaact	cctgggagtt	ctacactcag	caaacgtttg	ttgttggaact	1740
atgaaggcgg	acacagattt	tatacgaatt	tgtaatgcta	acatctagca	taagaattgg	1800
caaccataga	aaatactacg	tgtatatata	tgtttatagt	ctcaaaaaaa	aaaaaaaaaa	1860
aaaaanaaaa	aagggcggcc	gc				1882

<210> 77

<211> 2892

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (858)

<223> n equals a,t,g, or c

<400> 77

agactctgag	tccagctccg	aagaggaaga	ggaattcggg	tggttgga	atcgctctcg	60
ctttgccaa	ggagactatt	tacgatgctg	caagatctgt	tatcgctct	gtggttttgt	120
catccttgct	gcctgtgttg	tggcctgtgt	tggcttggtg	tggatgcagg	ttgctctcaa	180
ggaggatctg	gatgccctca	aggaaaaatt	tcgaacaatg	gaatctaata	agaaaagctc	240
attccaagaa	atccccaac	ttaatgaaga	actactcagc	aagcaaaaac	aacttgagaa	300
gattgaatct	ggagagatgg	gtttgaacaa	agtctggata	aacatcacag	aaatgaataa	360
gcagatttct	ctgttgactt	ctgcagtga	ccacctcaaa	gccaatgtta	agtcagctgc	420
agacttgatt	agcctgccta	ccactgtaga	gggacttcag	aagagtgtag	cttccattgg	480
cmatacttta	aacagcgctc	atcttgctgt	ggaagcacta	cagaaaactg	tggatgaaca	540
aaagaaaacg	atggaattac	tgcagagtga	tatgaatcag	cacttcttga	aggagactcc	600
tggaaagca	cagatcatc	cgtcaccttc	agccacatca	gaacttgaca	ataaaaccca	660
cagtgagaat	ttgaaacaga	tgggtgatag	atctgccact	ctgaaaagac	agtctttgga	720
tcaagtcacc	aacagaacag	atacagtaaa	aatccaaagc	ataaagaaa	aaggatagtt	780
ccaaattctc	cagggtatccc	aagcttaaga	gagraactcc	agcttgatcc	agtgtcttta	840
cmaacmaacc	tgrgagcnac	mggcctccag	agaccgccga	tgargagcaa	gtagagagtt	900
cacatcarag	ccatcagcat	tgccaaaatt	ttcacagttt	cttgagagcc	cagttgagaa	960
agctgcccac	ctaagacctc	tctccctacc	aggagtctct	agcactgaag	atcttcagga	1020
tttattccgc	aagactggcc	aggacgtgga	tgggaagctg	acctaccagg	aaatctggac	1080
ctccctaggt	tctgctatgc	cagaaccaga	gagcttgaga	gcatttgatt	ccgatggaga	1140
tggaaagata	tcattctctg	agctaagggt	agcttttaggt	atctagcttc	atcaggcata	1200
ttttagaaat	ggactgccta	atatctattt	acctaaacac	aaaacaaccc	ttacttacc	1260
atcagtcctc	tagtctctca	aactactgta	gcagatactt	tgccaccttt	taacttgttt	1320
gaagaagcta	tataaaagtt	atttttttaa	agaagaagac	cattttactt	atgatgttca	1380
gaaatctatg	atttctctca	accagtaaga	tcttacattt	taaaattgcc	agaaaaaaa	1440
ttaaagccct	ctttttttct	ctttcttttt	tttgagggga	ggagacctta	tcttttaaa	1500
ctgggaaatg	tatatagaga	gagaataagc	cacttttata	tttcacttaa	atttgcttta	1560
aattagctgc	actttataga	gactcagaaa	atgtcttttc	tttaaaagat	aggccttttc	1620
tggttgtaaa	tatttaaatg	aaagaaagca	ttgtgcatat	tggttgga	gtaggaagaa	1680
tggttttgaa	caggatatga	acaaatgact	tattaaaaat	tgctgatctg	gtgtaggtgg	1740
cagctgaaac	tacatccatg	tctccataag	gyatccctca	aaggcccagg	cgctgccagg	1800

gggtttgtcc	tggttagctgg	aggaaccgat	ttcagggagt	agacactgga	gacaatactg	1860
actccaggca	tggtcatg	aagtaggatt	ctggttcttt	gttcctattc	cctcagctaa	1920
tcccaacctg	ggaatcagag	aagtcttggg	gatttttctc	atttttagta	ctatttcagg	1980
gtttatgagc	ataaaaagt	atccattggg	gagctccatt	ttccctgctg	agttagctag	2040
attgccttcc	ccaccacccc	acttaagtct	gtcttaaagc	cgtagctggc	tcccaccacc	2100
agtaccatct	ccatttgaat	ggcagggcta	aattccccca	gccattatct	cacactgacc	2160
acccagagct	ttagaagaga	gctgtgcttc	taattttgac	ccagaaaacc	atacccttg	2220
agattttacc	tagaggctaa	ccaagagcct	aatatgttcc	tctgggggat	gactaaagcc	2280
aaaaaggctg	tgagatgaaa	catgtgaaat	aattttcagt	ttccttacca	ttaccagctc	2340
agaagtagct	agaggcttcc	tacccaaagg	atgccaaagt	atagcagggc	aggcctggag	2400
ctagggcctt	cacatggtgg	tagcaagttt	ttcaaatcta	atacaatcaa	gtacaatacc	2460
tccttttaaa	gcttctgtgg	acctggcatg	aaagatccct	agattgaaag	gaataatacc	2520
tccatgtctc	ctgtatgttg	agtctagaat	tgctgtgttg	ttcttagaag	cagtctttgg	2580
gcaacaactt	gaaaggggaa	aaaaaaacta	caaaaactta	actttggtat	aggccaagtc	2640
aggagaaaag	tagagaaaag	tgcatgcca	cagacttctt	tagtgagat	catttctctt	2700
ttacttttgt	tcaggttgcc	cttcaccatg	gatacagtc	ggtagccctta	aacattttaag	2760
ggctgttttt	tttttcttta	catgatgttc	agcttggtat	taaccaaact	taaatttttt	2820
ttccagaagt	attaaaattt	agttaaagca	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	2880
aagggcgccc	gc					2892

<210> 78
<211> 1673
<212> DNA
<213> Homo sapiens

<400> 78	
cggcacgagc	tggaatgaa atttgccctt gtttatatgt acctgtcttt tatttgcttc 60
tgtctttttt	attgcaactc aatagacaca caccattgct ttgtctctga ttatttgcca 120
tttgaatcgt	caatgagggg agctttttaca gaacttttga tactaataaa aggtgagtc 180
aatgttttaa	aaaagatgca gaatcatcat ttatgtcaga gctactgact cacacttaaa 240
ttgcagtggt	agcactgaaa aagaaatgta tatggatggg aatatagatt gcaggccaat 300
taggaccctt	cttttgaagt tgggaattgag ggatagctac tgttctcttc tatctttgag 360
ggttaggaga	actttattca gtgttgaata actgtattcc tctgtttat taatgtttgt 420
tgtggggggt	ttctattcag caccatcttc tgctgtctct gctccccgc cccagagga 480
ggataataa	agaggcatgg gacaggggct tataataata agacatggga ggggttgatt 540
acccagtgtc	ttcaagtaac ttttacgaga gatttgaaat agccagcgat caatgcaaaa 600
tagcaatggc	cttggcagaa ttgtcacata cataactcaat gtttacagtt taaactctgg 660
tgtcagacag	ggatcatagtt accccgattg gatgcattcc atctctgggt cagaacctct 720
aaaacttggt	aaatcattga aagtcattct cttattaaaa aagcagattc tcagactcac 780
atcagactag	gagaagtcct gagaaatcta aatttttagc acatgctttg ggggattctt 840
tacatcacgt	gtgtttggga aactgtgctg attgatgtcc atggaaaagca gcctcaggca 900
tggggagggg	ctggaaaaga attatttagg tcagtttcgg gatcttagat tgtttcttgg 960
ctacactggc	cactttttaa agtgtgctta gaaagagtat gacacctttt taattttcaa 1020
aaggacttgg	gttcagtgtg tgtccttatg tttaaagaac agccctcttt gtagttactc 1080
tagaaatagg	tagaatggca gaaagagcgc tggctgtctg tgttttaggc ctgttttgta 1140
cttcatgtgg	ccatgtggta tgggaacatc ctgggatttc tgtgagcctc tgtgaactca 1200
gattccccat	ctggaaaaca ggagtaacaa cactgtgtgg aacctttatg gagtgttaaa 1260
aaagtgatag	ctctttgtaa gcgacgaaga gccaggtcag tgtttaattt tattttctca 1320
gaaatagtac	tagttattaa ggcctttaac aaaaaaaaaa ctttgaaaag gctaattggg 1380
gcctgggtata	gtgtgtcatg cctgtaagcc cagcattttg gcaggctaaa ggggggagga 1440
tcacttgagg	ccaggagtgt gagagcagcg tgggtaacat ggtgacatcc tgtctgtara 1500
aaaaataaaa	acattagctg aatgtgtgtg catgcgcta tagtcccagc tactcggaag 1560
ctgaggtggg	aagattgttt gagcccagga gggtagggga agctataatt atgccactgt 1620
actccagcct	gggcgacaga gtgagatcct gtcttaaaaa aaaaaaaaaa aaa 1673

<210> 79
<211> 1461
<212> DNA
<213> Homo sapiens

<400> 79	
ccacgcgtcc	ggagagttat ggagaatgct gattttgatt attatgatgc cagatactga 60

```

gaatatctta catgtatctt ctgagcagag cttctgttcc acaaagttaa atccatgctt 120
aatataatctt ttgccaagta aatttttagtc gattgcacct cagttgttga ttagtaaccc 180
atcggcagta gaaagatggc agtgtttttt ccaggctgtt tggtctctta agtatctaga 240
cgaggccgag tcagccttat ggtctaaag ctgccaatct tctgtgtgtt tctttatttc 300
tttatccctt tatccagctg ctacttactg ctattgccac atttgccctc tggctcatgg 360
gatagcatgc ttagcttccc ctgaggctac tgtaaatgct tcttttttac tctgtgggt 420
ggaaatgtac ttggcatcct tagtcttaaa cctctctccc ctcttttttc cacagacacc 480
aggcacttaa gttagcattt cagcctgcac cagttatcag tagtagcttt caaccctca 540
tttctgggtc ggtaactcag cacactgtcc caagagagct tgactaagcc aatttgcccc 600
ctcttccctt ctccctctgt ctgttcatct ttcttttttc ttttctctac ccatccattt 660
ccttgactct ccttttattt ttctcttact ctctttaatc tcccaaatga tttttttctg 720
cttttagtat agcagatgcc ccagaattag gcagataact gtaatacaaa ataaaaaat 780
agtaaatctt aaaattaac atttgcctca gattggatca actaaaaaac gagtttattt 840
tttatgactg gtctattcgc ccctttatgg ctataatgca gattttttgt attaaagtgt 900
tataggtttg tggtttttgt ttttttgtgc ttttatacaa agagttgtga agatcggttt 960
tatgcaggcc tgctcattca agatgatctg tgatgtggga aaaaagttaa atctttttct 1020
agctaatggt ttacaaggaa aaggaaagct acttttattt ttatttattt atttttttac 1080
atacaatgat tcgaatacac agtttgagtt atttttcaaa ctaactttct ctgaatatgc 1140
tataaatggt ggctgttcat ttttcaagta atggtttgta aacaactttt aggcattctt 1200
agcttaactaa tttttatgac caatagttaa ggacataaag attataccta tgaattgggg 1260
gatcaagaac agtaacagtg ctctgcaggc ctcgatcatt aactgccaac aaaatctaca 1320
ggacaattcc aaatgtctgc aaaagaaaaa catgaaaaat tcatactgat aattatagat 1380
cagaatcatt taaagccctt atctccttcc tctctcatt tccctaattc taattcttct 1440
ctctggaaaa aaaaaaaaaa a 1461

```

<210> 80

<211> 1517

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1145)

<223> n equals a,t,g, or c

<400> 80

```

aggagaaact ctaaaaactg cagatattat ttcatgctat atgttccatc ctctgatgag 60
aatgtgagga aagaaaattg tatcctgcat ggtgaaaat ggtcccctac aaaaatatca 120
tgttggacaa ctaactctgag atagtggat ctctggaaag cagtttagca ctggtgagtt 180
tggactttca tggcaggctg ccttggttca tatcttttgg taatgatact tatcctctgt 240
ragggccatt tctttatttg tggaaatgaa gacaatagag tgcttagata taatttagaa 300
caatgtccgt cacatagtaa acacgtaata aacggtagct cttattgtta ttattattac 360
tattattacc ttgaagacag gggctctgtc ttgttcatca ttccatctcc agctcttagc 420
acagtccctg gcacaattca aacatgtatt tggatgaatg acaaatagct actgaatatt 480
tgccctgttc caagcattgt tagaggta ca tgggacaggg cagtgaacaa aacagacaaa 540
acctcctgct gtctcagagt tcacactcta atggggagac ccaggcaatg aggaaataat 600
taaaatatac aatgtgtctt atggcaataa atgacaaaaga aaaataaagc agaggtgaga 660
aacagtggca gtgttttggg gatcatttgc tttgcaacaa gccactcccc aaagttagtg 720
gcctaaaaca atttaatcac agttcatgtt ctggctacaa caatacacat cctctctatg 780
tgcaaaaata actcactcct ccctcagagc ctctgacct taagggttca ggttcaaagc 840
ttaagatctt atcctctgaa gtagggttag ggacaaaaca gtcttctcag gtacttcttc 900
tggggacaca gagacttgtg aactaaaaga caagttacct acctccaac acaactgaca 960
tgcaatgggg atataggaag agataatttc aataggcgtc tctgtgcaaa agcgggggaa 1020
atgagagta ctcagcagtc acggttcata ttaactctaa atctagccag gcatatatcc 1080
caagtcttcc tgatgtgagg acaagaatta ttcttgatt agggctcact twwtctctt 1140
gaggntgggt cgcctcagct tttggatttg tctctgaa catccttctc tgtctataaa 1200
atgcatgtat atactcatc atacatagag agaaagagag agagagagag agagagactc 1260
tgtcacgcag gctggagtg aatgggtgta tctcagctca ctgcaaccta caactcctgg 1320
gttcaagcaa ttctcctgtc tcagcctccc gagcacctgt agtccctgct actcaggagg 1380
ctgaggcagg agaattgctt gaatccgaga ggcagagggt gtcagtgagc agagattaca 1440
ccactgcact ccagcttggg tgacagagca aggccttcatc tcaaaaaaag acaaaaaaaa 1500
aaaaaaaaa ctcgtag 1517

```

<210> 81
 <211> 574
 <212> DNA
 <213> Homo sapiens

```

<400> 81
tagtagagcgc cgtgtataga ggcagagagg agtgaagtcc acagttcctc tcctccaaga      60
gectgccgac catgcccgcg ggcgtgccca tgtccaccta cctgaaaatg ttgcgagcca      120
gtctcctggc catgtgcgca ggggcagaag tgggtgcacag gtactaccga ccggacctga      180
caatacctga aattccacca aagcgtggag aactcaaaac ggagcttttg ggactgaaag      240
aaagaaaaaca caaacctcaa gttctcaac aggaggaact taaataacta tgccaagaat      300
tctgtgaaca atataagtct taaatatgta tttcttaatt tattgcatca aactacttgt      360
ccttaagcac ttagtctaata gctaactgca agaggagggtg ctgagtggtg gtttagccga      420
tacgttgaaa ttttaattacg gtttgattga tatttcttga aaactgcca agcacatatc      480
atcaaacatc ttcatgaata tggtttggaa gatgtttagt cttgaatata acgcgaaata      540
gaatatttgt aagtctacta taaaaaaaaa aaaa

```

<210> 82
 <211> 1455
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (390)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (456)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1100)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1293)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1409)
 <223> n equals a,t,g, or c

```

<400> 82
gggtccaccct cccccagggg cctccccagc ctccctctcc acctccctgc cccccggaga      60
tacctccaaa gccggtacgc ctgttcccag agttcgggtga gtgctgcagc caggagatgg      120
ggctctgggt ggatggcctg ggatccctgg aatcaggcct ctggaaggta tgcaaggatc      180
acactccttt ctgtgcaagc ctgccaccag cccactgtgt ggccccgggc aggtcacagc      240
ctccctgagc gctattctct tcactcctcac aatggagaca gcacccacct ctctggcctc      300
ctgaccgtta agtgtggggc catggccggc tttgccagtt acccatggtc tgattttcca      360
tgggtgtggg tgggttgctt ttctttttkn ttttttttt tgagacagag cgagagtctg      420
tctcaaaaaa aaagacaagt tgcagatgag ctgagntttg ggcagagcaa gcgggattct      480
gatggggggg ggatgttgcg ctgcgtcagca ggcaatagtt agttgggtga gggttttgat      540
camgggggtg ctactgcctg ccccatttta tccagctctg tagttgctat agagtgtgta      600
gaaccttggc acatcactta tcagttttgt cacctcagat ggcttcttca ctacttgggg      660
tgtctcctgg gtgtggggct ctccctcctg tggcctctgc tgactgcctg gcaactggcac      720
acatgctctg gtgagggggg gaccaacggt ttttcccggt tgtttctctg ttccctcggtt      780
aacctcctc gtcttgtaag atgaatgtwc ttgtctctgt tcactatgca gatgaggact      840

```



```

ttgaggctca gagacgccac taacttgccct ggtccaagcc ttttgggcct ctcaggctgc      900
agccagcaat gctgcagtga agtttgcctg ggaggctgac cctaggagtc tgcaggcgctg      960
ttaggacccc cgatctagaa gacagcagag atgtaggcca gggaggacca ataccgagca      1020
tctgagggca ggcacacctc agactgacca gaatacaaat gaattcgagt cacttacaaa      1080
caaagtggca taaggccagn cacagtggcc catgcctata atcccagcac tttcggaggc      1140
cgagggtggga ggattgcttg aggccaacga tgtgagacca gcctgggcaa catagcaaga      1200
ccttgtctct acaaaaaataa aaattcaaaa aagtggcatt taacacatac ttttttctt      1260
ttttttgaga cagarttttg ctctgtcccc cangctggag tgcaatgggt tgatctcggc      1320
tcaactgcaac ctccacctcc caggagaact gcttgaacct gggaggcggt tgcagtgagc      1380
caagatcgca ccacttcact ccagctgna caacggagca agactccatc taaaaaaaaa      1440
aaaaaaaaaa ctcga                                     1455

```

<210> 83
 <211> 1640
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (687)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (764)
 <223> n equals a,t,g, or c

```

<400> 83
gtgagcactg gtttaagcac ctcatagact ggcatttctg cctcccacaa gataactgga      60
cctggccttag gaacttgaat agcagcatgc atggagtgtc tctatgtgtc aagcactgtt      120
ttaagcacgt ttgaaataac tcacttcatt tgaataaact gagtctacat gatgactgta      180
aaaggttgggt tctgttggtta cctgcatttt accgatgagg aaactgaagc cttcagaagt      240
gcagtcactt gtccagggcc acatagcagg ctgagatttg aaccggcagg ccttttgact      300
ccagagctta cactcttaac tccattcatc tgctaagtcc ttccctgtcc tcttgcaaga      360
tgctttaatc cagggattat caaacttttt cttaaaatca ggagaactca ttgcaaacca      420
attcatacct agattccaca gaatcaaaga tgcagccgag ttaccatttg agctggagtg      480
ggggcggtara attgcctgtt ttggcctcct tcttgacatt gctgttctta ctgcagcctc      540
tgatgcttcc ccttgagggc tcccagacct agttgggcaa ccacagtgtt gtccgtctgc      600
ttctcccgat tcagaggctc ggcttttgccg aagtcctccc actcgaagtg gcacagagtt      660
gaggtctctt ccaggcacac tggcgtnccc tcactgggct cctgtccctg ccttggtcaa      720
catcctggtg cgcactgggt ggggtgactaa caacattttt gganttggtg ctggagccca      780
ggtgactact ccaaatcacg gttttccatt ctgtgtgaga tggcctcatg cctttctatg      840
cctctgacag gcagttctct gaatttcgaa ggctctgtc ttaagagact gtcagaagtc      900
ccttttgcaa gggactgtgg gcaaaccgcc cagcggctgt ggtcaattcc tctctctgat      960
ggcagtagtg ctacctaggg ggccgcctgg gtgaaacggg cttttttgca tacttccaaa      1020
ctggttccct gtacttaggg gaccaaacaa ttattgtctg aaccaagatg ctcttgagag      1080
tgaagagaat gtaaagtgtc cagtcctgga cagatggtat atatgatcgc cgtaaataca      1140
gccagccctt gccagaagtg ggtctggaga aatggtgcgg gggggcggtg aaagggttta      1200
caaccgcgag tctt tgtct ctgctagggt aattggtagc atcagtcctc actctgctta      1260
ttcagaccaa aaaattgtta agttcttccc accaccacgg agcacagact tgattaagat      1320
ccagaaaagg cagccgggtg cagtgaactg cgcctgtaat cccagcactt tgggaggccg      1380
aggcgggtgg ctcaactgag gtcaggaggt tgagaccagc ctggccaacc tggtaaaacc      1440
ctgtctctac taaaaataca aaaattakcc asgcatgggt gcccatgcca taatcccagc      1500
tactggcggg gctgaggcag gagaattgct tgaaccgggg aggcgaaggt tgcagtgagc      1560
tgagatcgtg ccatgcactc cagcctgggg gacagagtga gactctgtct caaaaaaaaaa      1620
aaaaaaaaaa aaaaactcga                                     1640

```

<210> 84
 <211> 525
 <212> DNA
 <213> Homo sapiens

```

<400> 84
ggcagcagga gaactgatgg ggggtggagag aagctccttg tgggaggaga gggaactacc      60
agcagagccc ctccctaccgc agacacagga tcggagacaa cctccaaccc cacctgcctc      120
ctgaagtgtg gctgacatgc aactgcctta actttgccta cctggcctcc ttatgatccc      180
cctccggcgt ggtatggttg gggggcttct tttgctgctg gccacggcaa acaagctgct      240
tgctgcttcc ttcagagacc tcatggatgt tcttacatgc ccccgacccc ggtagatggc      300
tcctgttgtt ttggggagcc tgggaaggtgg ttatgccttt tggatgcagg agaggagcaa      360
gaaagagtgg agagggagaa tgggggagcc ggaccctgac ctccctgggt tctgggttga      420
gatgaaaaaa ttagaagcat caggtctaag atcagcttct cttggaagca gagcctgaga      480
caagatataa atgccagtca ttattataaa aaaaaaaaaa aaaaaa      525

```

```

<210> 85
<211> 837
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> SITE
<222> (717)
<223> n equals a,t,g, or c

```

```

<400> 85
cactatagaa ggtacgcctg caggtaccgg tccggaattc ccgggtcgac ccacgcgtcc      60
gggtgggaga tgattggctc atggcggttg acgtccccct tgctggcttc gtgatagtga      120
gtgagcgctc atgggatctg gttgtttaga agcatgcagc acctcctgct tcaactctctc      180
tgtctctcct gctccaccat ggccagaaac gtgcctgctt ccccttcgcc ttctgccgtg      240
attgtcagtt tcytgaggsc tccccagcca tgcttcctgt acagcctgca raactgtgag      300
tcaattaaac ctcttttctt cataaattcc ccagtttcca gtagttcttt atagcagtgt      360
gaaaacagac taatggaccc ttctggttga aggaatgyag ccattctgct tgtttrasta      420
tktcctttct attcatctct atttccyggg aggtgtttat ccaagtgcaa taggagrtat      480
tggtgacygc asagtccctc cagtgttctg ctagtaaata gttgaagggt gatcaktgat      540
ctycwgcrtt ttcagtctgg catggaaaag ccccyrtgya actggtaaag rtatcartaa      600
gcaccaggag gtatctaaat ccaccaggag ccataggcat cacgttgacg tccatttacc      660
agtcttcctt ggcaagattc ttctgaattg tgctgccttg gccaaaagag gtatggnagg      720
ggctgggerc agtggctyry gcctgtralc ccagcaggag ttcgagacca ggcaggagaa      780
tcactagcag agaatatgtc tccccaaccc ctctcaaaaa aaaaaaaggg cggccgc      837

```

```

<210> 86
<211> 1574
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> SITE
<222> (19)
<223> n equals a,t,g, or c

```

```

<220>
<221> SITE
<222> (873)
<223> n equals a,t,g, or c

```

```

<400> 86
gtgatctttg taatatctnc tgttgtttct atgatatagg agctagggga agggggttgt      60
ttgccttctt caggacctga ctggacagat ggacctggct caagcaacta ctctggatgc      120
actttgctgt gtgggatgaa ctaaaagtgt ctgaattttg ctgataactt tataaaactc      180
actatggcat gcttcctctc tgggtggccc taggatggat gacactcaag atactacaga      240
tgtgggtgca gccatgcaca cacacgatgg aatatggcca ttcctacaca ggtggggtag      300
agagtgggtc agcagcctgg cacctcacag aggtgggacc taagaggact catgattatg      360
cagagaattg gattgggtct ctgtcataga ttgagtaatc tcttccttca cctcaattcc      420
atctccaccc atctctacat ctgggcacag caaccagag atggccaaaa gcattcaagc      480
ctgggggaag atgtttgact attgtgctc ttcaccagaa cctcacacct ctctgggac      540

```

tggaaccctt	cagtgggtgt	gtggccagtt	ttggaggctg	gaatgatggg	ccaggggtga	600
ggattcattc	tccatgtaaa	gtttcctttc	atcctgccta	gccatcccca	aggtttattt	660
ccagaagaaa	ggaatatctc	tacttggatc	aattctggtc	atttcaagag	gatggaggcc	720
tcaagtgtgg	gaacttcccc	tactccctgg	atgtgtgtac	ctagcacact	tccttctccc	780
accccttttt	ccagtgggat	ttgtttttct	gttctcttct	gtcctgtctt	atactgcaac	840
ttgtctctct	aggggacaga	tggtcttctt	tgncatcttc	actctccacc	cccagagagg	900
agtcagagcc	ataactcaat	cactcagccc	ctccaaagat	agttgatgtg	tgataaatctc	960
ataatgttga	gaaccctgat	gagatacatt	gtcttctctt	ccctacaatg	cctctggggc	1020
caaggcacc	attcttcttg	ctatcctcca	tccccctga	ggcttccact	tttttttttt	1080
ttagacataa	agctgggcat	cagcaactgg	cctgtgttga	tgcaaagctg	ctttgctctg	1140
tatctggctg	gactgatctg	tctcacaaga	agccatgagg	ccataggagg	aagctccctc	1200
tccccctcat	cttctgctcc	aaagggtgta	gcaagaggag	tacccagtta	gggggttgag	1260
cccccatata	acatcttctt	gtcagaagac	tgatggatct	ttttcattcc	aacctctccc	1320
ctttcccccg	atgaatgcaa	taaaactctg	tgacaccagc	aaccattgct	ctttagaaat	1380
gggttttctg	atcatatggc	tgatgtgtta	tgggcagtat	ggatgtcttc	attgtgtgct	1440
tctgtttttc	atcttttttg	ttttattaat	aaaaatttat	gtatttgctc	ctgttactat	1500
aataatacag	ggaataaatt	attcaatcca	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1560
aaaaaaaaact	cgag					1574

<210> 87
 <211> 1628
 <212> DNA
 <213> Homo sapiens

<400> 87						
tcctctcttc	ctgcctgtgc	tcaaacaccc	acagagaact	ctggaaggag	aagaaaaaca	60
tacctctctc	cttccccggc	ctttccccac	acacactgag	gttgagaagc	tggaataaaa	120
cagcggtccaa	cactttttaa	tggcgggtgg	agctccagac	aagggaagaag	ggaaggactg	180
agagaagaaa	ttattagact	tttttagactg	gctgaaccca	gaacctttac	attggttaca	240
aaatcatgcc	tggaggtgaa	aaattcaata	gcattatacc	aagtgtttgg	gcaacgggtgt	300
agagacaata	gtagagacaa	gctacagctt	tagacagtgg	tagagagaag	ctacaacttt	360
agcttgtaga	aacagctgtt	gggtgttcag	cagatgcagt	ctaggtgcat	gcaaaccac	420
tggtgtgcagc	atgccttcac	tctcacgaca	aggagagccag	accttgtgtc	tgcggtcggc	480
tgaatattgc	atggaatctg	tggattcaca	gaggcttctt	ctcagctaag	agggagtgtg	540
gctggatgca	tttctctgtg	gctcattcaa	tttggggtat	actcatacta	ctgagtctct	600
atgaaggagt	gataagctgg	gtcttcaatt	tccaaatgtt	taccaaactc	ctactatgtg	660
ccaagcacta	ttcccactgc	tttgagagct	aacggtaaac	aacaacagyg	aaaaaacatg	720
taacatccca	accagtgcct	ctaaggattt	aaaatatgct	ggtacttatg	actttattct	780
tacttctgta	ttatagatat	gtatatggct	ttggggatag	tgtatatgta	cacatatatg	840
cacacatata	cacacacaca	catatatata	atcagctgtc	catagcctac	tcattctctga	900
taattttatat	tttgtactca	aattttctga	ataccccac	caaatacttc	tctcctcctt	960
accaatatta	taattgtcct	gacaataaca	taacaaaccc	agctctaaca	cctacagatt	1020
tctttgagaa	caaacaactt	ctacatgcaa	tttcttttct	atactcacca	actggttttc	1080
ttcaacctcc	tgccccacct	gtccagctca	ggacatcaac	aaccttttat	ggaaaccacc	1140
gaggtcagac	tggatgcagt	cagttggact	gattcatcat	gactagtcca	attaagagct	1200
gatcaccttc	aaacagctct	gactttggaa	gcaattgatt	tgactgcctc	tttggtcaca	1260
tggccagatt	taccataaat	tttttgcaaa	cttggatgca	tcttttagata	cagagcaatg	1320
ctttggcatc	tgggggaggg	ggttgttcct	ggtgctgtca	cttgtaccca	ctctcctgtt	1380
tcctctccaa	ctatgagtca	cattttccct	caatatctcc	tatcttactt	tttgagtgat	1440
cagccctgac	tttcaagtct	aaattttctc	tccacgccaa	aaacaaaaca	aaaacagaaa	1500
aacaaaaaaa	gctttttgct	gtatcacacc	acctaaagtt	tggttagtga	acatgagcag	1560
acctcttctg	aatccccacac	atcagccatg	ctcttgcagc	catgtagagg	agctggagggt	1620
gggtgggc						1628

<210> 88
 <211> 1795
 <212> DNA
 <213> Homo sapiens

<400> 88						
ggcacgagaa	caaactataa	actacttacc	tgcattatgc	tttactggga	aaaatcttag	60
cagatgatac	ttcctttacat	ttgtagagta	gaatgtgttt	tatgtctttt	attagtatat	120

atgactggcc	ctatatcatc	taatagatag	tccttttcat	catggagatg	aattattgtg	180
ggttccagagt	tttgtatatg	tctctaatec	tgctagggag	tccaatcata	cccttgtggg	240
cctatacttc	agccacacag	gctgcagctt	tagtgacatc	acacgtgtgg	aaacctcttc	300
tagaggctca	ccagatcaat	atttctcctg	aaccttcaat	acattatgat	agatggcaca	360
ctcagagtaa	ttgtagttaa	ataaattctc	ttcaataaat	ggttctggaa	aaacaatatc	420
tatatgcaga	agaatagaag	aagactgccc	acttctaaca	atatacaaaa	atcaaatgaa	480
aattaaagaa	ttaaatctaa	gaccttgagc	tatgaagcta	ctacaagaaa	actttgggaa	540
aaatcttcag	gacattgacc	tgggcaaaga	ttttttgagt	aatactccat	aagtacaggc	600
aaccaaagca	aaaaatgaac	aaatgggatc	acatcaagtt	aaaaagcttc	cacacaacaa	660
agaaaacaat	caaagtgaag	agacaaccca	cagaatggga	gaaaatattt	gcaaaactacc	720
catttgaatg	ggattaataa	tcagaatata	tgaggagctc	aaacaactct	atagaaaaaa	780
atataataat	ctgatcaaaa	aatgggcaaa	agatttgagt	agacattcct	caaaagaaga	840
catgcaaatg	gtaaacaaac	atattgcgaa	gtactcaaca	tcactgatca	tcagagaaat	900
gcagatcaaa	aactacaatg	agatatcatc	tcattctcaat	taaaatggct	tctttttcca	960
aaagagaggc	aataactaat	gctgggtgaga	atgggaagaa	aaaaagaatc	ctcatgcact	1020
gttgggtggga	atataaacta	gtaaaaccac	tatggagaac	agtttggagt	ttcctcaaaa	1080
aactaaaaat	ggagctacta	tataatccag	caatttcacg	cctgggtata	tacccaaaag	1140
aaaataaatc	catgtatcaa	agaaatattt	gcactttcat	atttgttgta	gcaatgttca	1200
caatagtcaa	gatttggag	caacctgagt	ccacaaacag	ataaatgaat	aaagaaaatg	1260
tactatacac	aatggagtta	ctattcagcc	atgaaaaaga	atgagatgct	atcatttgca	1320
acaacataga	tggaaactgga	agtcatttgg	tttaagtga	taagccagaa	acagaaagac	1380
aaacatcaca	tgtcctcact	tatttgtggg	atctaaaaat	cagaacactt	gaactcatgg	1440
acatagagag	tagaaggatg	attaccagag	gctgggaagg	gtagtgggag	gaagggtggg	1500
gttgggggtga	aggatgtggg	gatgggtta	gggtaccaaa	aattgaatga	ataaggccta	1560
ctatttgata	gcacaacagg	ctgactacag	ccaataatag	tttaactaca	ttttaaaata	1620
actaagagta	taattggatt	gtttgttaaca	caaagataaa	tgcttgaggg	gatggatgtc	1680
tcattttcca	ttatgtgatt	attacacatt	gcattgcctat	ataaaacatc	tcatatctca	1740
tgcaccccat	caatataaac	acctactgtg	tatccacaaa	aaaaaaaaaa	aaaaa	1795

<210> 89

<211> 1864

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1844)

<223> n equals a,t,g, or c

<400> 89

cccaagccag	ccatttatta	caagaagcaa	caggttattg	acattacatg	tttgaaaatt	60
ccctttgggc	tttagggaaa	ataaacagga	agccaagatt	tggagccttt	gtaataagga	120
cttctctgag	aaagtctttt	ctttactata	attgagtaat	tcatatttag	agtcacatgt	180
ccagtagcat	ttctaatttt	gagcattcac	cttgctacct	ttaaaaaaca	tctgagtttt	240
aagtggcctt	tttatcatca	tacacatgtg	catacaagaa	agggacttgg	cagtttaaaa	300
gccacatata	ttcactttta	ttgccctaaa	tttacctgaa	acagacatac	tggcaaacctc	360
acatattgct	ggtgctaacc	ttatatttca	tagtgttggc	atattccctt	tttcttagat	420
tcttactccg	aaatataggt	acacatcctt	tgtctgtgtc	agagggaatt	acatcctttt	480
tcctctccta	caaaaaacatg	ctttatttaag	tatccatcat	tactttcctt	tatgctcgct	540
caatatgcaa	tgtgctgtaa	ttctaccatg	taccttaaat	aaaggatgat	ggcaaaagtta	600
tttaccatgt	agaaaccatt	ttctttctag	aaacaatagc	tcagcctcac	tgtagcagct	660
ggcatgtgtg	gtcaagtggg	tagttgtact	cttgcaagtt	ggatttaata	tcatatatac	720
tggaccttca	gactgttaaa	aatcaatgta	accttttttt	attgctatgg	caagcaatta	780
gtattttcact	gcacgtcttc	catactaagt	ttcatttcta	aatcttatat	gtaggcattt	840
gttagttcca	atgatttctt	cactaatata	acacttttta	atgggaatct	ttccacctac	900
agccctggaa	tgataatgct	acagtaattc	ttctgaattg	actttttctt	tcatectgtc	960
agcttttggc	aatatcccaa	ttatggcagg	gaacagggtg	ggaactaaga	tcagttacaa	1020
aaagtgttag	atgtgtcaac	tttgtatggc	tgggatcact	gtgcccaaac	aaaacaggcg	1080
aaatacctca	gttaaaattt	ttccatcaaa	gtcttttaaa	gaagagtata	ctgaagaaag	1140
ggcagtcata	atacttactt	ctaacagctt	ctaaagggtg	catgtttaac	atttcatttc	1200
aaaatcacc	caaatttgca	ctaaatacca	atgaagtgtt	attttgcctt	agtagtcttc	1260
tgagcaacaa	actatgggga	attctgkaaa	amcatataaa	aagtycaagm	cttttttttt	1320
aatgaatga	ttactatgtt	aatgcaaac	tttttttttt	ttttatttaa	acaaacatac	1380

acttctctctg	gcaagggttat	agatgattaa	cctctgttca	tagacttata	tataaaacta	1440
gaggggttttt	tgtttacttt	tttaattttt	caagtgcgat	tgtttcttac	acagacatta	1500
ttactattaa	attatcattt	agccagttat	ctgcaaata	atagtatgta	ttgtctcttc	1560
ttgtgacgtt	tagtttaatt	gcttatttta	aagcagaaam	attagttaca	agtgtcttac	1620
aatattttta	ccaacagtaa	agtagagact	taatgaaaat	accttagtgt	gattttaata	1680
taatttgcat	atttttagttg	tataaagttt	taattgtaaaa	tgccattat	tgaaggga	1740
agatctttca	ataaaaaata	cccacgaaaa	aaaaaaaaaa	aaaaagggcg	gccgctctag	1800
aggatccaag	cttacgtacg	cgtgcatgcy	acgtcatagc	tctnaaaagg	ggactccaga	1860
gctt						1864

<210> 90

<211> 1983

<212> DNA

<213> Homo sapiens

<400> 90

gacgttgaag	atgagaacaa	gcagaagaaa	caattggatt	tctatgaaaa	gaaaacagat	60
tggtgtacac	ttacacaaat	ttgtgcagat	tatttgtcta	gaaggaaagt	catacagggt	120
gggcagtcctg	gtcacaaaaa	gggacagggg	ttgagggggg	tctggtgact	gtgatgaagg	180
cctcactctc	aggcctccgg	tcccactgaa	ggtcagatga	aaggtagtct	tccctggcgg	240
ttgctgctgc	cactgaatgg	gccctaactt	tgctgtcttg	tgtttgaatc	ttctgcagga	300
cacgttagcg	tatgccacag	ctttgttgaa	tgaaaaagag	caatcaggaa	gcagtaatgg	360
gtcggagagt	agtcctgcc	atgagaacgg	agacaggcat	ctacagcagg	tataacggtc	420
agcatgtcct	tggttgcaaa	gggcagcctt	gctcttaagc	tttccaaaaa	gaatttccac	480
agctgaggga	aaacaagatg	cttctctctg	aatgtgagtc	caaagagtta	ccagcgctgc	540
cctctagtga	tctcagctca	gcataatcac	taaccgtgtg	tttacagggc	tgagttagtgc	600
tgcatgttga	agtgaatgga	aggcctcgag	gtgtttgttg	ctggccaccc	tgatcagcct	660
gcaggtagtc	ccgatgaagc	cagggcacag	ggggattcgt	tccagcttgt	tcactttatt	720
ctgccttgcc	aggttactga	aagtcctctg	tttgctctca	ccagccttcc	tggaatgtg	780
gactcttgaa	agaaaaagctc	ccgtgctctt	gaagtatacc	tgcttgccag	gggagtccaa	840
gaaaattttg	acatgtattt	ttaaaaaaag	aaaaaaaaaa	agctttaata	ccaatcatta	900
tagtagaaaa	agaaaaataaa	tatgtattga	acccccactg	tgtagcaaaa	ctgaactaag	960
tgtagttaa	tcattacgtc	tttccaatag	tctgtaactt	tccttaacag	cagtcctctc	1020
tgtaggtcct	tcacagtact	tggtacagaa	taggccccat	taaatgaatg	ttactgatgt	1080
agtaggtgtc	attttttttt	aagtgttatc	tttcggatcc	tcataagcac	tatgtgaggc	1140
agctgtcacc	ctgattttac	agaaaaggtaa	ctgcagccca	gcacagtgat	gtgacttagc	1200
ccaaggtcac	tccacacatt	acctcatcac	ctacttcatt	tgtagagaaa	ataaaagctg	1260
tcacaggaga	gtcctctcgg	ccactaatcc	ccaagcatct	gcactgttct	tgtstcctct	1320
cctgtgacag	tggaagttt	gcctctgtcc	acccaaagcc	cctagcgctc	atccccgcc	1380
accttggcag	agctttgcgt	tctaattgtg	atgtaactct	tcaatatcca	gaacgctyca	1440
ccctgccaga	cccttcccag	cgacgtctca	gcacactggg	ttctctctctg	ccctgtcaaa	1500
gcctctcttc	tgccctgtca	aagcctctct	tctccctgtt	gccccctgct	tcttttctct	1560
tctttgcagc	caaacttcga	ctaattctct	aaacttaact	ttccccattt	tcttatctct	1620
caactcgtct	tcagcctctt	ccctgctaac	tccctcttct	ctccaactca	gcagttgggg	1680
tgacaggtgg	cctgcagctt	tcaggcctca	tcttagccga	ctgctcggca	gcacttagcg	1740
ctcctggcgc	tcttcccgtt	tgaaacacta	ttccagggct	ttcctgacac	ttctctctcg	1800
tagtttttct	caaaccttcc	tggtgttccc	ttctctgtct	ccttcttagt	actgcctctt	1860
ctggaccacc	agtaaagggt	tgtaggtctc	ctaacctgta	tctctctgcc	ctcactccat	1920
attctctccg	cgtcgacgcy	gccgcgaatt	cccgggtcga	cgagctcact	agtcggcggc	1980
cgc						1983

<210> 91

<211> 1957

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (349)

<223> n equals a,t,g, or c

<400> 91

```

gtttctctgc aattatttgg ttcagtttta attttatctg taacaatgat agtaattgct      60
gtttcacttt ctctctctctg tgatgttggt tctcttgaat gtatgagctg ctttactcct      120
aagttttgctg acattgtttgc aaatgcttat cagaatgaat cctatatattt tatttaaaat      180
gatcgtgtca ttttcaatca ggcagcccat ccaaacatgc ggacctatta tttctgact      240
gatacaggaa aggaaatgga gttgtggatg aaagccatgt tagatgctgc cctagtacag      300
acagaacctg tgaagaggta aaggcttgta gaaaaaatga tgggtattnc cacttccatt      360
ttatttccatg ccttgcaagt atttactgt catagtmcat atcattttta tagtcatggt      420
atgaaatcat ttttctctca gaaagcaagg atcaattcct gttctgaatt aaattaatac      480
acattttgtt agtttgcat accacctaca tttttattcc acttttcttc tttttcttct      540
ttattcattt tcacctatcg gtgtactggg gtgaatccag aatcctaaca ttcaactga      600
atgttcttctc ttcttacaga attaccttta atttccggtg agtacgtttt taattgttac      660
cttaaaagcta cacagatttt tatcctttga gatagtgttt ttaagattct aaatcttaga      720
agagagttta tttttatgaa gtttaattgt gtttttctgt aatagtgtgt tgatgtttct      780
aaagtgtgat gaattacagt aagaamcttt gatartttca ttttttcaac atttctgatt      840
aatttttatt gtttttgtaa tgaatgtctc cagaaaaatag ttcgtcaagc atttagattg      900
tttccaaatc cacttcttgg tgaattgtac cttttttata ttgaaactcc actactcaga      960
tyccttgata ataatagataa gtgctgttaa aattgaccca tgtatttttc cctgcttgaa      1020
gatacgaatc attttaatat tcttcagtat agctagttag aggaaatctg atttctcagac      1080
tacataaata caagtagtat aatgtgcttt ttaaaaaata tgtattcctg taattcgaag      1140
aaaaaattat gatgcaagtt aatttttctt ccagtcagtg acagctgagc acatatctta      1200
tgtaagaaag atgctaattgt gcatcttttt tccctcttct ttttttccc tcttctcagg      1260
aaattaggga ttgttmcagt atacatctag tcttttgggt ttcttattct agtgtgcatt      1320
ttaataaagt cttggctttt tggctaaaag acttaggttg atgctgtgta ttgtgtctat      1380
ttttgtaaat atcaagtcta aatcaagtta ccaatcact agtaattaga gctggggaaa      1440
aactgaaaag aaaagagggt ctaggatata gctctaggac atctattttt aagaaaaacc      1500
acttttgcca catgcatatt gcaggatgag agcagataga aggaaatct gtttttgaa      1560
ttgcatgtgt aaaaattacc tgagtgcatt aaagtgcatt tgggttagcac tgataacgag      1620
agaaaatgtg taggtgaaga gaattcattt aaaatcttca ggctgagcat ggtggctcac      1680
acctgtaatt cttagcattt gggagggtga gggatcactt gagcccagga ttttgagatc      1740
agcctgggca acatgatgaa acaccatctg taccaaaaat acaaaaaatta gctgggcgta      1800
gtaccacact cttgtagtcc ctgctactcg ggaggctgaa gcatgaggat cacttgaatc      1860
aggaggttga ggctgctgtg agctgtgact gtgccactgc actccagcct ggacaacgga      1920
gtgagaccct gtcacaaaaa aaaaaaaaaa aactcga      1957

```

<210> 92
 <211> 573
 <212> DNA
 <213> Homo sapiens

```

<400> 92
ggcagcagtg aatattaact gtgttatttt tatacacttt ttaagcctta actcgccatt      60
gatttaccag ttttaacgttt cctgggggttt ctttggccat ggggttctct gccccaccac      120
ccggcccttt gtttgacttg cgtcgtctga tactcagtat tgtagctttt tgtccgcag      180
ttactccctg taaatacgtt gttatacata ctgttaacac ccttttgcct tttctatggg      240
acctccaggc caccatattt agaactagtt acctatttaa aaaagaaaaa acagtctgtt      300
ggcttctcag tctgcatctt ggaggcaggg aggtgagggc aggtgcccct cagacacttc      360
aggaaggtag tttgcattct atttaaaaaa gggagtgggg agcaaatgaa aatcaaatgt      420
gggggggaaa cactaaaggg ggcaagaaac aaaggaatta caaacctctt gctctttgta      480
tttctctgtt gtgaagaata aactgtacct gcaccgggaa aaaaaaaaaa aaaaaaaaaa      540
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa      573

```

<210> 93
 <211> 1212
 <212> DNA
 <213> Homo sapiens

```

<400> 93
gccctcccca gctagaatga acattctgcc aggcagggac caatatgttt tatctctgaa      60
tctctaggac agtgctaact aaataataga ccatcagtaa tattcgaata catgtatgga      120
taagggattt gttccaggtc acatagctag ttgttgctaa tagaaaggac aagtatgtag      180
ataccagcca cagttttttt agtatctcta cgccctattg ctttcccttt aactttaagc      240
ttggtcttac ccatattttc tgtagtttaa cttgtctttt gatccctcta aggggctgtt      300

```

ttatataaac	tcattgatcat	tgtttctttt	tctctctctt	tccttccctt	ccttctcttc	360
cttctctctt	cttctctacat	ctgtctcttt	ttcttctctt	cccagctctc	ctctctcttc	420
ttttttcact	tgtaggcttc	tgtaatttaa	tcaatatggt	acttattaag	cactgagtea	480
aatgtctaac	actgtactgt	atcctatgag	aaatgaaata	gaagcagatt	gaagacatac	540
cattacttga	ggaatttaaa	atctttattag	ccccttcttc	tcaatggcct	ttgtgctctt	600
ctgggtctgg	ttatctgtgt	tcttttctgg	ccttctgcct	tgaccatttc	ttttggcccc	660
tgctttggaa	attagtagat	aattttaccct	cattttggct	tcacatgac	cagctacagc	720
aagacccaaa	taagaaaaga	tgttacagcg	acattgatga	agttgggtcta	acacagaaac	780
tgaagagtg	agagagacag	aagaagaag	catgaagtag	ggaatgagga	gtagagaatg	840
tcaccaacgg	ggaattacat	gtgaccaaaa	aatcaaaaaga	ttatgactgg	gtacatatga	900
aaaataggtg	caggccaggt	gtagtggctc	acacctgtaa	tcccagcact	tggggaagcc	960
gargtggtg	gattgcttga	gcccaggagt	ttgagaccag	cctgggcaac	atggtgaaac	1020
cccatctcta	caaaaaatmc	aaaaattagc	csggcattgt	ggcacacaac	tgtagtctca	1080
gctactcagg	aagctgaggt	gggaagayca	ttgagcccag	raggcaraag	ttgcagttag	1140
ctgtgatcct	gccactgsac	tccagcctgg	gtgacagggc	aagaccctgt	ytmaaaaaaa	1200
aaaaaaaaaa	aa					1212

<210> 94
 <211> 1144
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (849)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (865)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1087)
 <223> n equals a,t,g, or c

<400> 94						
aattcggcac	gagggacagt	cagctaacta	ggcaagtcac	aatcttatat	agcataatca	60
tggagagtaac	actccatcat	ctttgctgtg	ttctattggt	tagaagcaag	tcactaggct	120
agcccatact	actgggagag	gattacacaa	gaacatgtgg	gtagaaatgg	gaataacttc	180
agctgtccaa	caatcttaca	ggtatatact	tcataatca	ttagctataa	gtaatatgg	240
gtttccatta	gtcaaaagatc	tgtgtgtcag	caagccagga	cttcaatatt	ttttaagat	300
ggtctttcta	gagaaaaata	cagtaataat	gggatgacag	aaggccatgt	gttttgtttt	360
gctttgtgtt	gtgtcttgg	tttctctct	atgactttgc	ttgttaycag	cttagaaaaa	420
actaacgcag	gtggggtgat	agcatggggc	tgtatctcag	tctctgtgca	gacacaaact	480
ttttctctc	ctaccagtta	ccaaacattg	tttattgcct	gtaagctctg	gaatcccaga	540
aaactttagt	tttaattctt	atcatcatca	ttatcacata	attacatcc	tagtttagat	600
ttggagcttg	tttttagatta	atackttaca	gagtagtttt	acatgaataa	gcttaaacat	660
tttccccga	tttttagttct	ctggcttacc	agaaaaatga	aaaacaacaa	caacaaaatc	720
cccaaaaactg	agaaccagg	aatgatagac	aacaaaactg	tgttttaatt	ttcatgattc	780
tagttgttca	acctgttttt	ttgacactct	gtatctgcat	tcattttatc	actaaaaaga	840
tgcttagtga	attgtaagta	tcatnttagg	cactgtgaat	tcattgataa	gatattctct	900
ctctctctct	ttttttcttt	tgagatggag	tctctgtctg	ttgccagggc	tggagtacag	960
tggcattgatc	tcgtcggctc	actgcagcct	ctgcctcccg	ggttcaatcc	attctcctgc	1020
ctcagctact	ccagcggctg	aggcagaaga	attgcttgaa	cctgggcagc	ggaggttgca	1080
gtgagcnaag	attacgccac	tgcactccag	tctttctcaa	aaaaaaaaaa	aaaaaaaaat	1140
cga						1144

<210> 95
 <211> 1274
 <212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (722)

<223> n equals a,t,g, or c

<400> 95

agctgagtg	gcgagcgcca	gggggtccag	ctgcacgtcc	caggctctcc	agcgcgcggc	60
aggccggggc	gggacgagga	gagctgcggg	gacaacgcct	gtggctgggt	ccggagtgcg	120
gggtgcggcg	gggacaagcg	ggcagcatgc	tcagggcggg	cgggagccta	ctgcgccttg	180
gccgcgggct	aacagtcgcg	tgccggcccc	gggcgcctct	cgaggccacg	cgacggcccc	240
caccggctct	tcgcggcccg	gggtctccct	gctactccag	cggcggggcc	cccagcaatt	300
ctggggcccca	aggtcacggg	gagattcacc	gagtcctcac	gcagcgagg	ccttcgcagt	360
tcgacaagaa	aatcctgctg	tggacagggc	gtttcaaatc	gatggaggag	atccgcctc	420
ggatcccgcc	agaaatgata	gacaccgcaa	gaaacaaagc	tcgagtgaag	gcttgttaca	480
taatgattgg	actcacaatt	atcgctgct	ttgctgtgat	agtgtcagcc	aaaagggtcg	540
tagaacgaca	tgaatcctta	acaagtggga	acttggcaaa	gaaagctaag	tgscgtgaag	600
aagctgcatt	ggctgcacag	gctaaagcta	atgatattct	aagtgcacaa	gtgttcacct	660
gaataccatc	cctgtcatca	gcaacagtag	aagatgggaa	aaatagaata	tttaccacaa	720
tntctgccat	ggttttat	tggttaacaag	aagcacaatg	tcttttttat	ttttattttt	780
tagtaaaatt	ttactgaagt	ataccatgca	ttcaaaaagt	ggacaaaact	gtatacagtc	840
tgatagatat	ttatgtcgtg	aacacctgtg	taaccactgc	caaagtgaag	atgtagaata	900
ttggcaacac	ttcacagcct	cattcctgcc	ttttctcagc	cattacctcc	caaacatagc	960
agtttttctg	agtttcatca	cctttgat	atgttgcctg	tttttgaa	ttataataat	1020
ggatttatac	attatgcact	tgtgtgtgtg	gattatttac	ctgacagtta	taagggtta	1080
ccacaaattg	tgtgtaccat	tagttcatcc	attgtcattg	ctgtattctg	ttgtataaac	1140
ataccacaat	ttattttgat	atgttgcaca	gtttctggcc	actacatata	atgtctaaat	1200
gagcacattg	tatatgtcat	taaaatgagg	ttgaactaaa	aaaaaaaaaa	aaaaaaaaaa	1260
aaaaaaaaact	cgag					1274

<210> 96

<211> 1780

<212> DNA

<213> Homo sapiens

<400> 96

tatttgggat	tatactgaac	ctatttgtcc	aataacctga	gttttcaaat	aatttttagtt	60
ctataagtag	tataattata	tataattata	tgaattcaga	ttagctgaaa	ggaaaaaaag	120
tagaagcctg	actacttggg	gctaactact	aaagattttg	gcagaatcaa	tggttgattt	180
ggcttctctg	tccttctccc	atgccagccc	cccagagtgt	tctgccttgt	gctgcctccc	240
ttcacckgga	gtgccacacc	cctctctctg	ccagttcagc	tcttctattc	tcaaggcctg	300
accttgtctg	acccttgtgc	ctctaaaccc	gtggccccac	ctctcttggg	cacgagctat	360
gtcaggtgat	gtttgtgttt	ttgggtatgc	ccatctccat	agccagacca	agcactctgg	420
aagccagggt	tgggtgctta	tttatctgtt	tgccatgcag	aaaatatctt	gcacaaaatt	480
acctctgtta	aggaatctga	agctgaattt	agtttggctg	agtcagggtt	gggttttttt	540
taaggggctg	tggggtgaaa	tgttgactgg	aagccaccca	caaacacaca	cctgctgggt	600
aggaaccccg	ctgtgggtgg	ttctgagctg	tttggcttca	ttgacagttt	ctgattgccc	660
tgagcaccag	gtctcatctt	gcattctctc	ctggcctgga	gaacattcag	tttctctcca	720
accttctcca	cctttctccc	actccttggg	aggaactgaa	gttgggggtg	aggagagcca	780
gatggctgga	gtgggtat	gaaggtcttt	ctgtcacctg	ttcagtggtg	tctgccccac	840
ccctgctgac	caagactgac	tgaatgttaa	aataatacag	accatctcaa	ctcagaaagc	900
tggcacattt	ttgaaagccc	aagtgtgggt	aagtgcgtgg	aacaacgata	attcacactg	960
ctttatgata	agaaattgtg	agaaatattg	tgccaggcaa	tttgcaaaat	cttgggaagg	1020
tgtgtgcact	taaccaccca	gcaactactc	ctggatgcat	cctagagaag	tgccatgtga	1080
acagagaatg	attttaagac	ttcactgaag	tattgtttag	gtagcaagat	tggaagaaag	1140
ctgcatttca	tcagcagaag	aatggataaa	taaatgagtt	gtttttgggc	cttggaaagt	1200
gaatatgaaa	gagttacgtg	tcaacacaga	tagatgaaaa	attatgtcta	gaaagtgtgt	1260
gaagctacat	acaaggtacc	cttagtgtaa	agttaagcat	actgtgtacc	tgtgggcacg	1320
ttacttcaac	ttgtttttca	ctttttctgt	aaaatgggat	agtagtgcca	atctcacagg	1380
gtgattgtgg	gtgggggggg	ggtcaatgaa	gtaatgcata	taaaatgctt	agaatagtgt	1440
ctagcatgta	agccttgtgg	acatatagaa	agtggtattg	ttttgcacag	taatctattt	1500
tctgtggatt	caaataatat	gaaatgagta	taaaatcatg	tattggaacg	atgtgtgcaa	1560

gtcaccattc	tgccttccta	aggcaggaga	cctgatggat	ttgggggggg	tacatggggc	1620
cttcagttgt	gttttctttg	tttttttcta	aaaattgatg	cagaggcatc	acaatgttaa	1680
gattttaaca	gggtagtggtg	gtgggtactt	tttaactggt	tgcttaaagt	gtttcaagat	1740
aaaaatattt	cttaaaaaaa	aaaaaaaaaa	aaaaaaaaaa			1780

<210> 97
<211> 2065
<212> DNA
<213> Homo sapiens

<400> 97						
ggcacgagat	taaaaggcct	ttcaaaagaa	tgggtttgaa	aaactcagta	ccctttaata	60
catgtacatt	tctttccttt	tttcatttaa	tgtaacatgt	ctgttgtaac	tatgtttctt	120
aaatattatt	ttaaggttat	gtgttcttta	attatggtca	aataataatt	ggtcacaaaa	180
aatgaaataa	tagtttaaaa	caagtagctg	ttactaagtg	tgctaaaaat	actcatttta	240
taatttaatt	tagttttctt	agtatattat	tataaattgt	gccctaagtc	aggtaacaa	300
gtacacatca	aaatgcccat	attgtatcta	tctgtagtcg	tttaattgtg	attatatgtg	360
aatttttttc	aaaattttac	taaccagaat	tctgtttatg	gcacctaacc	acgcagcatg	420
aggaaaacgg	cacaacacaa	tcttgagggtg	ccttctgaat	catcagatta	aattatgctt	480
cataatgttt	tgcttttact	gtatttcttt	aaaaactcta	aatctttatt	catgtgtcac	540
tggattaatt	tatctgataa	tgtgtctcac	aagaatctgt	tagatcgttt	attcttcagt	600
tgtactttga	atgggtgggt	ggaagtttca	ggtgaacaat	ggataacaaa	aagcaagtta	660
tggagattga	tgaagaggat	ggaaaaactg	aatacaagat	acccaaaaatg	aaaaaaagtg	720
tcccatTTTT	aataactata	ttctattatt	ttataaatgt	gtaataaagg	ggcccccttt	780
tattgggtct	tatccccctta	atcttttggtc	tttttcagta	attttaagtt	ttctgggatt	840
tttttttggt	tataaaactt	gtgttttagc	tttatcttgc	tatggagttt	tcacacttct	900
atagcacata	tcctagtatc	tagtcatttc	tgttttaata	tgaatttcag	taatttaatt	960
ttaatctggt	gacatattaa	tcgaaaataa	ggagtaattg	atacctccac	atgtcctttc	1020
tttttgctct	ctcttaaaat	cacaatatcc	agtaggagtg	gttatccaat	ttcttcgttg	1080
ttttaatcat	caaatgaagt	tagagaagta	tactaatccc	agcaactatg	actcatctag	1140
gcattgttaag	accataaaag	aattcaggaa	actattttcc	tgatttttaa	ataactttta	1200
gtgttatgta	acatctatcc	ttctgtttta	gacatgcatt	tcacatatag	ttgaattcta	1260
gattctaaga	taattcattt	tgggttaatac	ttcagagtac	tggatctaga	atcaggcttc	1320
ctgaatttaa	actcaggctc	cccatttaact	gtgtgtctgt	gagcccagtt	tctcatctgt	1380
aaaaatggggc	aacagtggca	ctcatcttaa	aggggtggat	aataaaaaaa	tgcatgtaag	1440
gccctaagca	tagtgcctgg	cacagaatta	ctgctcaaat	gttagctgtc	gtattaatat	1500
tgcaactttg	cacactgatg	tacatttctc	gttgaccagg	ctcattcttt	aagcattctc	1560
catgcttaaa	ccagttccat	aatccctagg	cctgtactcc	agggattgag	actgaaagga	1620
tcattttatgc	catgtttctc	taaaagcatc	attgctggaa	gacttttgat	aagtctgatg	1680
tgctcacaagc	tattctcagg	ccttttttgt	agagtttaga	aatgaagtat	ttgaatcaat	1740
ttagtatctc	ctttactatg	tttctctctt	taatctcagc	caaccccccta	cctgcaggta	1800
aaccagcatc	tcattaaagag	ctgggttggg	gtactctatt	ctgtatgcat	cataatagct	1860
taacattatt	tagtagctgt	aacttacagg	tttaatgcta	gatgaggatg	tctcaagccg	1920
tgagtgtgct	tgtgtaaaaa	tgggtggcaac	atcatctcgt	tggtaggaat	tttttacttg	1980
aattgttatt	ttgggaaaaa	gttaacagat	ttcttctgga	taaagaaaat	aaattggatg	2040
atgtataaaa	aaaaaaaaaa	aaaaa				2065

<210> 98
<211> 1154
<212> DNA
<213> Homo sapiens

<400> 98						
ggcacgaggt	gccgtgtgtg	tgtgctgtgt	taagtgtgca	tgtgcataca	tgtgcatgtc	60
tgtagggtga	cacatctgtg	tgtgtgtgtg	catgtgtgtg	ctgcatgtct	gtggggaggt	120
gtctctcgtg	agagcgtgtg	acagctggga	tttgactctt	tgctgtctgc	cccagagacc	180
acagccctgg	caggccctga	ccttctgtgc	ccgtgcatc	gagccgggtc	gctgcgggtg	240
cctgtggccg	ccaatgggga	actcgggtga	gctggcagga	gggtgtgccc	agagccctgg	300
ctgctgtctc	tgccactcag	cacagctggg	ccaggctgtt	gccccagagg	gcgtcagacg	360
tgaacttttg	gaacatcttt	attctgtttt	aaagttagca	caaattatta	gacactttcc	420
ccaaaatcca	tgtgtttggg	gcgtcttccg	gccatgccac	acatctgtgt	ttgcttggtt	480
gtttctgcac	cgagttccgt	ccacagcccg	ggtttctgtt	gttttaagtc	ttgagccctg	540

ggccgggggc	cacttctcat	tggtggctgg	aggctcggcc	aagtgagggg	ctgcttctgg	600
ttggagaggg	gagtttctgg	aaggggggtc	cccatgtgtc	tccagcgctt	cctgcagtct	660
ggggaggggc	ttggcaggag	caggtctggg	gagaaagccc	tggccggggg	tggaggtca	720
gtcctgggag	tgggcggggc	agctgggctc	gggggtgtaa	caggggtcctg	cggggggact	780
ctgtgctgag	tcaaaggagc	cggaagctgg	tgtgggccc	gtgggggtggg	gaaggtgggt	840
gcaggcaggg	gagggggctt	ggactgaagg	tgagaccag	gcctgggcaa	ggatgcgggtg	900
tgccagagc	ttggcagagc	atctgcctga	agcctgactg	tggcctgggt	ggggtaagga	960
agggttggag	aggctttggg	gcctgcggga	aagggggctg	tggagagaga	ggctgaccga	1020
gggctgccga	gaggaagacc	agtgttgctg	gagcctgtgg	tggagagggg	cttgggtgggt	1080
gaaccctcca	gggaaggcct	ggggcagggc	tcagaggacc	tgggaaggtg	gcagagttgt	1140
gtccagcagg	agct					1154

<210> 99

<211> 615

<212> DNA

<213> Homo sapiens

<400> 99

ccagggagac	agcagcgtgg	tcagagtggg	aggagctggc	catcggtgag	agctgctcca	60
tgcttggtcg	ctgggtgcta	gagcttggg	accactggct	tgcttactg	tggttggtgg	120
tggcggtgac	agagtgtgca	gcacgaccag	agtggctttt	ctggctttgc	ccgcccagct	180
gtcccatgcc	aggaggagga	ggagacacct	agagcctgcg	acaccatggc	tcgsctcgct	240
gcagtgtagg	ktctacccat	gtaacagatg	aggaaaccaa	ggagcacagt	tatttactaa	300
ctcgcacaa	gttcgaggcc	gagctcagac	ctgtggagca	gaagctgagt	gcgctgcagt	360
ccccgctggc	ccagargccc	ttcttcgagg	tgccctcacc	scgtgggcgc	gtggacctgt	420
acgagtaygc	atgcggggay	gaggacctgg	agccgctgtg	acgccgcccg	cgagaaacgc	480
cgcrccgggc	cgctccccac	gtgccaccac	cgggccaccg	cggctcgtgt	aaaaactgtt	540
gtggaaaaatg	agtgcgtttg	tacggaatga	taaaacttta	tttattcaca	aaaaaaaaaa	600
aaaaaaaaaa	aattc					615

<210> 100

<211> 1624

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (117)

<223> n equals a,t,g, or c

<400> 100

atatgtttct	gaatagatcc	agttgaatag	tctcattcaa	tttgagactg	ttgaacaact	60
gttgttttct	cacatacatt	taaagtcagg	gcacatgcgt	cactgcttat	ttttcgnact	120
tgacatatcc	cctgcatttc	catgtctgcc	tgtcctttag	ataaacagta	aaagtttccc	180
atgtgccagt	atttctcaaa	tggtttacat	cagaatcacc	tggaggactt	ttaagctaga	240
ttgtttggtc	ccatcccagg	gattctgaat	caacagggtt	gcgatgggtc	cagatagttt	300
actttttcaa	gtttactaac	aagtttgcca	agttccccca	gtttattacc	attagaccat	360
acctttttgt	ccaatcattt	aaamcaaatt	tttatataat	aagttttatt	tgtatgtaat	420
aaatttttatt	atataaaaaat	aagttttta	atatattata	taaaaagttt	taataaaatac	480
ctaataatatt	atttaatatg	ataaaactta	tattaaatga	aatttttatgc	tgtyctcttg	540
caaatctgtc	ttttgttatc	ttgctgggtg	gcctgtcatg	tgagggactg	caatctgata	600
tgcttatttt	ccacagtcaa	agcaattaca	agagaattgt	tacaattacc	cagttatgtc	660
aagagatttt	tttttaattc	actaaggtag	agataaggag	aatgtattaa	aataggatat	720
tttaattata	aatgcatgac	tggggagggg	gtattgtttt	tgaataaaat	atgaggttat	780
ttgccatgac	aaaaaaaaaa	agaagtagga	aaatcccatg	gaaatttatg	ttccttctaa	840
cttttaaaac	tacctaataa	atataattga	tttaaatat	atctcaatat	tccccattct	900
tttatatccc	cttaaatagg	tacccatgaa	gagattatga	actacttgaa	ggtggagact	960
gtacgggtgt	gtgttgagc	tggcttga	tgtcttatga	gtgacaatcg	ttagtttgag	1020
gaattttgtg	agacagtgtg	caaattgttg	ctagcttgaa	atctgcggca	attggagtat	1080
ttacaccata	gaaatgctat	aagtgaagrc	ctacctttcc	cttaagagct	agttgtttaa	1140
cttttaccag	cataccactg	gaccttgtct	aaaatttctt	tgtgttccca	gtgtcttgcc	1200
cagtagatac	aagataaata	ttgccagaat	cagatatcag	gaagtagtaa	gaaaaggagt	1260

taatatgcaa	actaaatcac	tcgctcaatt	gaataattga	gatcttctgt	tcatttggtc	1320
cttggacctt	aatcatttgc	atthttggaga	aaattttttc	tgcttttaaa	gtctgtaatt	1380
tcagtttttg	tgctggggag	agggaaaaac	tatttgcctg	tagttgcttt	ttgtgacaaa	1440
gtgaataccc	actgggctaa	gtttcatatc	taaagcttgt	cactaagaat	tttcattttt	1500
aggggtcaca	aacctatttt	gaaaatagtg	ttgtgtgaat	gctgtaagtg	ttgtacatgt	1560
ctctgggttc	agaattaaaa	gaattcagag	ttaaaaaaa	aaaaaaaaa	aaaaaaaaa	1620
cgta						1624

<210> 101
 <211> 1756
 <212> DNA
 <213> Homo sapiens

<400> 101						
ggcagcaggtt	tccctctcac	atatatattg	ttttgtgtcc	ctggctaaag	tacaagcttt	60
ttgaaggcag	aaaccatgtc	tttggtttct	tttgtatttc	ccatagcacc	ttttactgtg	120
agagtgggca	cacagtatat	gttgtggaat	gacatcctga	gtgatccctc	cctggctggg	180
cctcagatta	aattccctga	aatggaacag	tcctaacca	cacaggacag	gtattctcca	240
tctggcatgt	tggttgctcc	tttcaacctg	ctatttgaaa	tggtccctt	caacatcttt	300
ctgttcccc	agtggggctt	gctatggcta	atgctgtact	tgctgtatgt	gttccaggcg	360
agtctcggga	caccagaact	gacctgggag	cgagtggat	ctcaagttga	ccagtgtat	420
ggcctgatgg	caagaggata	gtactgctgg	cagaggtaag	ctgagactgg	caaaaatact	480
ccccacacac	aggagagact	gcaataccca	ggtccctccc	tcctcatgtt	ctcgaatact	540
ttcaactcct	ctgttaagca	caagtttgac	tactttccca	atggatttta	cttctaattg	600
tgaaagatct	tttcattcag	caattaagaa	actattttgg	ttccccactt	ttcaccaatt	660
atcctgtctc	tccacgtcaa	tccacaggtt	gagtttagata	attattacta	tagaaggaat	720
tcacagatag	aaccagtgcc	actttgagtg	atgcatacaa	agagataatg	tcacttgttg	780
gatgttttaa	tcactaagca	caaagtagat	atgcccgaat	gtaaccagga	ctatcttagg	840
caagttctgg	gaatgtatgt	ttttactgat	agattccctg	tttttgaaat	ccattccctt	900
gaattgagcc	agatgagtat	aggtacctac	ctagatatca	attgtcfaat	tgatatttcc	960
ccatcctagc	tcctagctca	cattgacact	attgactttc	atthttattg	cttccatgtc	1020
agtgtttgac	cacttttctt	ttcttaaaag	ctcctcttcc	ctagtccctg	attcctgaca	1080
gctataatat	tagatgcctt	ctattcttac	cttgaagctt	tctctctctc	agagaaagat	1140
acaaaaatat	caaggaggat	aataatactt	ttctcaattt	tgattttcag	ttgtgttttt	1200
ttcttttttt	atattaaaga	acctgaatat	gaaaatgtaa	aatatacatt	gtcttttatc	1260
aggggcccac	aagttaggag	tttttagtgt	ccttactgtt	tcttcacatt	ttcctcactt	1320
tatctcatct	tctcagatag	ttcagggtac	ttgtaaaggg	actgaactat	ttcttcacaa	1380
ggaaggagta	tatatgagga	ggagatgggc	agattgccaa	atatgcatta	atagctttga	1440
tgctagcttg	ctgactgatg	acttgtttct	agctgcccta	ggagggtcca	cctggtaatt	1500
ttggtgacaa	aagcaagtac	catgggtgtt	tttggctaga	tggttgagca	aaaagggtgg	1560
caggcttcat	aggaacacaa	ataggaaagg	gtggcattgg	gggcaatttc	tagttcttct	1620
actgtctgaa	tcaccaactc	aaaatacaag	gctgacaatg	ctgtctttga	attcaggaga	1680
agcaaaactga	aggagaagca	caaaaatcat	cacagctatg	gtgaaacctt	gtctctacaa	1740
aaaaaaaaaa	aaaaaa					1756

<210> 102
 <211> 1416
 <212> DNA
 <213> Homo sapiens

<400> 102						
tacatagtta	ttctttttta	ttttttactc	aagttacatt	taatatcttt	atcacaggaa	60
ggctggcaat	ataaaacttc	ctatgtacga	aaactcaaaa	ataaccaaag	tggaagtgta	120
ataattcctt	tgagaagcaa	aagaacagta	caatgtttat	taacacgttt	cttctgaaat	180
tttcttcaat	ttttttaaac	acacaaaaag	ctttttctga	cttagattgc	tggtttgctgt	240
ttttaatggt	gttaacatgc	atattattgc	atttatggat	agtagtagat	agtgtaatat	300
acatgaaacc	aacatctagg	gatgggtgcc	ttctgagtgc	tttacagatg	gcacgtttct	360
ttattatcca	gcttaatcac	agctcctcca	actgataact	tcacatcctc	tgcatgtatt	420
ccaatctgta	aacttggttg	gcacaagttg	gttttagcgt	atthtgaacc	gtatttttaa	480
tcactggaac	tactttgcct	taatgcccat	gggctgtcag	ctccaagggt	ctaagaccaa	540
gtttttctta	actttgtgca	tatagcgtgg	gacctgccca	gaacaggtac	tcaacaattt	600
tgctgagcag	aactgtcttc	aatggagaaa	agaaaggaga	aaggctttac	tgaagactgc	660

cccaaataca	aacaaattcc	attttaaatg	gaatatatac	actttagccc	ccaaatgcag	720
accagtgac	gtctgtgtag	tttccgacta	gtcacctggt	aatagatcat	tcctgtcatt	780
cacaggctca	gtcccagctc	tatttttcag	tatcttgaat	caagttctct	ctcctcta	840
catggaagaa	atagacccat	aactagttat	tttgggtaaa	tgggagctat	ttaaaaaatt	900
gatattttaa	aagcttaaat	gaaatgttaa	tcaaataatga	tttatgatta	ttttctttct	960
atgagtattc	tttaattgtg	gaaggcagtt	tcttaggaag	ggaacaaggg	ttctctttta	1020
caacccaaaag	tttggtggtg	gttttttttt	cacaaaatta	ttgagttaa	aaaaattgat	1080
ggttggtttg	catttcacct	agtagcttat	tcaatgggtt	gtttttctgc	taaatgttaa	1140
ccgtcaaaac	ttgaattaat	ttcttaattg	ctatttctac	ttcagggaatc	ttaagaaaga	1200
tggcttaacc	cagtcagaag	ggaca_gcat	aattttcttc	catggctatt	taagtaagta	1260
ttaggagagc	tttcacgacc	atgctatagc	ttcctagtga	cgcagaattg	gtaagacttg	1320
tgtgatatat	acatgtgtga	cttgttacat	atcatagcaa	ctgtgtagtg	ggaaggatgt	1380
aaacagttcg	atatcaagct	tatcgatacc	gtcgac			1416

<210> 103
 <211> 704
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (287)
 <223> n equals a,t,g, or c

<400> 103						
actgtgtctg	tctgtctct	gatatttata	tgccattatg	tggcctctac	tgccttagga	60
ttctaagtgt	cccactaaga	tcagctaact	cagttccact	acagtgttta	ccaccatcat	120
ctctcgcaaa	caaagacagc	cacttcagag	ctcctaggaa	atagtgggtg	tcccatcatc	180
attgcattcc	tttaatsacat	gggtgaaaatt	aacaatggct	aaggagcctt	tgtgttttct	240
cctctacaat	atgcccagga	atttctggca	ttttggccat	cttattnata	ggctattact	300
gaatttmagc	ctmatcctmc	caaattatta	atgccaaaat	attaactctt	gattcttagg	360
tgagtgcacc	catgccaaata	aatttgccat	gatctaacct	taaatgtatt	ctcatatatg	420
ctgtccaagt	ttctrtgat	taaaaatggca	aggcctttag	ttctcctaca	taggttttct	480
ctctccagag	aaggcctcaa	ttctctgact	aggctatggt	gggatataac	tggaggcact	540
aataggtagt	agggtaaat	ctttatttta	ttatttttgg	agacaggagg	ggctctgctt	600
tgttcagact	ggagtgcagt	gggtgtgatca	tggctcattg	caactttgaa	ctcctgggag	660
acagagcaag	actccatctc	aaaaaaaaaa	aaaaaaaaac	tcga		704

<210> 104
 <211> 1259
 <212> DNA
 <213> Homo sapiens

<400> 104						
gacggggacc	agagcaggtt	cctggctgca	gaggccacaa	gtcacgctgt	ctctgagagc	60
cacgggtggc	tcattctctc	gccataaact	tgccaattat	cctgctgctg	cctcattgac	120
ttcgcaacca	ctcttccctc	tggaaacagag	gacactctcg	ccagctctcc	ccatggcgga	180
tccttgctcta	gggtcaggcc	tctgtcccaa	agtc-cccct	ggggacacct	tctctgacca	240
gccccctcatt	cctatggcct	catgctgttt	ttatttcttc	ctaggactta	gcacgtatcc	300
tagaaatttaa	cctgctggta	tatectgttt	cttgtctgtc	tctttccagt	ggaatgtcac	360
catcgcccag	gtggggattt	ttgtgtgttt	tgttccactgc	tgtacamcca	gccccagca	420
cagcgsctgt	ccaggacaag	tgcccagtaa	acacttgagg	agcaatgcaa	gcgtgcgtgc	480
atggataagt	awttctttss	cagatgaggg	ggctaagggt	cagagaaggc	cctgggggtc	540
tcagactcat	agcccagtcg	tctttctgct	gacacgcctc	ggtctctggg	gcagtttggt	600
gcctgttcag	caacaaagag	gggtgtgcctc	gttaggggtc	ctgcgtgcga	atcgagctcc	660
ctgcgtgtct	tggctggagg	tmacmacctc	ytctgtctca	gggcctgtaa	ttaccactta	720
ccctgggtcaa	tgggtccgag	agattmccct	tgtaggcagg	gctgtggcca	gggtgctcac	780
ctggccccca	gsaggtccca	tgggcactgt	ctggccgggc	ttcatggctg	acattccagg	840
tacatttcta	gcctgggct	gccatgggca	gaggggtggg	agaggggtcg	gggcttcagg	900
ctggacaaac	cagtcagcct	tcccagctgg	gcgcctgac	cacccaactc	ctgtggggct	960
ccttgaggcc	tggagggtgg	aggggggtctc	tgttcaacc	ccacccatgc	cctcttccct	1020
tctctccctc	ggcaggtcyt	cccagcagyt	cctgcaaaca	gacccccgac	ccaagccctt	1080

ccttctgsct	ccactgccac	cactgctgct	catctctgct	ggcacagaag	tctcttccct	1140
ggctctccag	aaatccccc	tccacactca	gccagagggg	gctattaaaa	ctgtgggcca	1200
gccacatca	gtccacagca	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaactcgag	1259

<210> 105
 <211> 1804
 <212> DNA
 <213> Homo sapiens

<400> 105						
ggcagagcag	acgcgkctcc	ttggaggagg	tgcggtcctc	tagggaggca	tcgggctcct	60
aggggcttct	tggcggtgtg	ggtgggattg	gggtccgcgc	gccatggcct	tcactttcgc	120
tgcgttctgc	tacatgctgt	ctctggtgct	gtgcgctgcg	ctcatcttct	tcgccatctg	180
gcacataatt	gcctttgatg	agttaaggac	agattttaag	agccccatag	accagtgcga	240
tcctgttcat	gcgaggggaa	ggttgaggaa	catcgagcgc	atctgcttcc	ttctgcgaaa	300
gctgggtgctg	ccagaatact	ccatccatag	cctcttctgc	attatgttcc	tgtgtgcgca	360
agagtggctc	acgctggggc	tgaatgtccc	tctacttttc	tatcacttct	ggaggatatt	420
ccactgtcca	gcagatagct	cagaactagc	ctacgacca	ccggtgggtca	tgaatgccga	480
cactttgagt	tactgtcaga	aggaggcckg	gtgtaagctg	gccttctatc	tcctctcctt	540
cttctactac	ctttactgca	tgatctacac	tttagtgagc	tcttaacgca	aagaccatgc	600
acatcatcag	agactgagat	gggagaggcc	tgagacggag	aggtgcattt	ctgctggtga	660
ctggaggagg	gaccagaatg	aggatacgtg	agaaatagac	ccggcaggca	gtcagactga	720
atggggagctg	gaatcacgca	gcagctggga	gccgagttaa	ccctgcgtgt	ctgtgtcacc	780
ctgtttgtca	atctttggca	ttcgaattcc	acacacgggg	tcctagagcc	cttctgagca	840
tcagtgtgtg	gggggagtag	gtgacgaaac	actagacctc	tcctgagaga	gaattgctgc	900
ttcctgaatc	cacttcattg	aacagcacct	tgcaagttca	aatgagttcc	tgggagtggg	960
ggctggaagg	ccacaagggt	cttgctaagg	aacagaatga	ccagagtgca	aggccaagtc	1020
tgcagggacc	tgttgaagac	ctcgagaatg	kcttggtctg	ccaagactct	tgktgccttt	1080
cttccaagcc	atggccatgc	cctttttctc	aaatgggarg	ggctggargg	tgtgtgggat	1140
ttgtcttcag	ctgcaaccag	ccttgagcct	gctgggctat	tttcagctga	ggagggtgga	1200
atataggaaa	aatgcatttt	tgaaacrttt	gcaacatgat	caaggtgtta	gttctccacc	1260
acacaagttg	tattcttctt	ttgccacctc	aaaccatcac	agagtcttta	aatgcaaatc	1320
aattgggtcaa	tgctagtcaa	agctatgttc	ttacaaaaac	cccagacagc	tcagagctca	1380
gaaaatcctg	tggagtggct	gctctgtacc	gtgggcatcc	ggcagccagg	aagttagaca	1440
acataattat	aactttgttt	tatgatgctg	catcatttgt	actgtttagg	tcgactgag	1500
gacatcatct	tatttagaat	tttccgtttg	gcattctctt	ttgggtggga	gttatgctgg	1560
gggttgtaaa	taatgacaag	gctgagattt	ttatgatgtt	taaattgggc	acaatgattt	1620
tgaccttatt	ccccaaactt	cttttctttt	ctactgttta	acatacacag	gctatttata	1680
cacgtcccca	gctcccatct	gaaacctgtg	actcagggtt	atgaatgggtg	tttgtgtagc	1740
aaacacatgt	gtgctatgtt	tattaaaatg	cagcgacaaa	aaaaaraaaa	aaaaaaaaact	1800
cgag						1804

<210> 106
 <211> 971
 <212> DNA
 <213> Homo sapiens

<400> 106						
ctagccccgg	cggatcccc	gggctgcagg	cgccgaggct	ggaggccgag	ctctgcagag	60
ttacaattga	gactgctaac	ccctaccttt	gaagggatca	acggattgtt	gttgaacaa	120
catttagttc	agaatccagt	cgactctctg	caacttttag	gtggtacttt	ctattttaac	180
acctcaagg	tgaagcagaa	gaataaggag	aaggataagt	cgaaggggaa	ggcgcctgaa	240
gaggacgaak	aggagaggag	acgccgtgag	cgggacgacc	agatgtaccg	agagcggctg	300
cgcaccttgc	tggctatcgc	ggttgtcatg	agcctcctga	atgctctcag	caccagcgga	360
ggcagcattt	cctggaacga	ctttgtccac	gagatgctgg	ccaaggcgca	ggtgcagcgc	420
gtccagggtg	tgcttgagag	cgacgtgggt	gaagtctacc	tgaccctcgg	agccgtgggt	480
tttggggcgc	ctcggctagc	cttgatgtac	cgaatgcagg	ttgcaaatat	tgacaagttt	540
gaagagaagc	ttcgagcagc	tgaagatgag	ctgaatatcg	aggccaagga	caggatccca	600
gtttcctaca	agcgaacagg	attctttggg	aaatgccctg	tactctgtgg	ggatgacggy	660
agtgggcctg	gccatcctgt	ggtatgtttt	ccgtctggcc	gggatgactg	gaggcaccgc	720
cggcgatgga	cgtccaggtc	ccggctcctg	tgctggaaag	cgttgatggg	gagcgtcggc	780
gctgaccaca	ckcgggagct	gcggaagccc	agcggttcac	acaggcctcc	cttcaacgta	840

gtcatccct	ggtggtgaa	gcaagacgac	ggccccctgac	gtgcagccac	acacagaaaa	900
ggctgctgtg	aaacatttta	atgcttcgac	tttttttttc	ttccagcctg	gagcaacaag	960
agcaaaactc	c					971

<210> 107
 <211> 821
 <212> DNA
 <213> Homo sapiens

<400> 107						
gttttgagtg	tgtgaattac	atatatgaac	atctgaraaa	atcctataag	cagtttaatc	60
aactgttcca	ctccactcca	agtgaagcca	taggcagaat	tgagttatgg	ggagagcggc	120
ctagtaataa	ttgggttgcg	taatacaaaag	ttctactggg	tagtgatgtt	gtagaagttc	180
atatagaatc	agctgagctt	tcagaaatgg	tgaaggggtg	gtaatagtca	taacttagat	240
tgtaattttt	ttcccatagg	cttttaaaaa	atattcatga	ggttcttttt	ttatttcaat	300
agtttttggg	gaacaggtgg	tttttggtta	catgataagt	tcttcagtg	tgattttctga	360
gattttgggt	cacctgtcat	gtgagcagta	tgaactctac	tttatgtgta	gtcttatccc	420
tcattgtgat	gaactccacc	ttatgtgtag	tcttatccct	caccactccc	tgcccttccc	480
cacaagtccc	caaagtccat	tatatgatct	ttatgccttt	acatcttcac	agtttagctc	540
tcacacaact	tattataatt	tataagtaag	ccagcattgg	atatagttgt	attccattat	600
taatttaaga	aaccttatgc	aagtaattat	tagtcatcat	ccccaaaaaa	agggagaaca	660
gggttagatt	cagaataact	tgataagagc	taaatactat	catgagtgct	gtcagtcgtt	720
agtaactttc	cattggtatt	ctatgtcttt	taggcttaca	gatacttttt	acactcttac	780
aaaatgtgca	caagaagaag	ctgcagctca	gagctcgtgc	c		821

<210> 108
 <211> 1779
 <212> DNA
 <213> Homo sapiens

<400> 108						
aggaatacat	acgatccttg	tctaccagga	gtctaataga	aagatggaca	gcgtggaccc	60
tgccagcagc	caggccatgg	agctctctga	tgccaccctc	attgaggggtg	tgggtaatga	120
ggtgatgggtg	gtggcaggtg	tggtggtgct	gattctagcc	ttgggtcctag	cttggctctc	180
tacctacgta	gcagacagcg	gtagcaacca	gctcctgggc	gctattgtgt	cagcaggcga	240
cacatccgtc	ctccacctgg	ggcatgtgga	ccacctgggtg	gcaggccaag	gcaaccgccga	300
gccaactgaa	ctcccccatc	catcagagga	caagcaggtg	caggcagcag	cagtccagag	360
gccccctga	gatctgagga	tagcacctgc	ctccctccca	gccctggcct	catcaactgtg	420
cggctcaaat	tcctcaatga	taccgaggag	ctggctgtgg	ctaggccaga	ggataaccgtg	480
ggtgcctga	agagcaaat	cttccctgga	caagaaaagcc	agatgaaact	gatctaccag	540
ggcgcctgc	tacaagagcc	agcccgacac	ctgcgttctc	tgaacattac	cgacaactgt	600
gtgattcaact	gccaccgctc	acccccaggg	tcagctgttc	caggcccttc	agcctccttg	660
gccccctcgg	ccactgagcc	acccagcctt	ggtgtcaatg	tgggcagcct	catggtgcct	720
gtctttgtgg	tgctgttggg	tggtgtctgg	tacttccgaa	tcaattaccg	ccaattcttc	780
acagcacctg	ccactgtctc	cctgggtgga	gtcacctgtc	tcttcagctt	cctagtattt	840
gggatgtatg	gacgataagg	acataggaag	aaaatgaaag	gcatggctct	tctcctttat	900
ggcctcccca	cttttccctg	ccagagctgg	gccccaaaggg	cggggagggga	gggggtgaaa	960
ggatgtgatg	gaaatctcct	ccataggaca	caggaggcga	gtatgcggcc	tccccttctc	1020
atccacagga	gtacagatgt	ccctcccgtg	cgagcacac	tcaggtagaa	atgaggatgt	1080
catcttccct	cacttttagg	gtcctctgaa	ggagtccaac	gctgctggcc	aagctcagtg	1140
gggagcctgg	gctctgagat	tcctccccac	ctgtggttct	gactcttccc	agtgtcctgc	1200
atgtctgccc	ccagcaccga	gggctgcctg	caagggcagc	tcagcatggc	cccagcacaa	1260
ctccgtaggg	agcctggagt	atccttccat	ttctcagcca	aatactcatc	ttttgagact	1320
gaaatcacac	tgccgggaat	gaagattgtg	ccagccttct	cttatgggca	cctagccgcc	1380
ttcaccttct	tcctctaccc	cttagcagga	ataggggtgc	ctcccttctt	tcaaagcact	1440
ttgcttgcat	tttattttat	ttttttaaga	gtccttcata	gagctcagtc	aggaagggga	1500
ttggggcaca	agccaagccc	ccagcattgg	gagcggccag	gccacagctg	ctgctcccgt	1560
agtcctcagg	ctgtaagcaa	gagacagcac	tgcccttggg	ccagcgtcct	accctgcccc	1620
actccaagga	ctgggtatgg	atygtgggc	cctaggctct	tgcttctggg	gctattggag	1680
ggtcagtgct	tgtgactgaa	taaaagtcca	ttttgtggtg	aaaaaaaaaa	aaaaaaaaaa	1740
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	agggcggcc			1779

<210> 109
 <211> 1518
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1146)
 <223> n equals a,t,g, or c

<400> 109
 aggagaaact ctaaaaaactg cagatattat ttcattgctat atgttccatc ctctgatgag 60
 aatgtgagga aagaaaattg tatcctgcat ggctgaaaat ggtcccctac aaaaaatca 120
 tgttgagcaa ctaactctgag atagtggat ctctggaaag cagtttagca ctggtgagtt 180
 tggactttca tggcaggctg ccttgggttca tatcttttgg taatgatact tatectctgt 240
 raggccatt tctttatttg tggaaatgaa gacaatagag tgcttagata taatttasca 300
 acaatgtccg tcacatagta aacacgtaat aaacggtagc tcttattgtt attattatta 360
 ctattattac ctggaagaca ggggctctgt cttgttcac attccatctc cagctcttag 420
 cacagtcctt ggcacaattc aaacatgtat ttggatgaat gacaaatagc tactgaatat 480
 ttgccctgtt ccaagcattg ttagaggtag atgggacagg gcagtgaaca aaacagacaa 540
 aacctcctgc tgtctcagag ttcacactct aatggggaga cccaggcaat gaggaaataa 600
 ttaaaatata caatgtgtct tatggcaata aatgacaaag aaaaaataag cagaggtgag 660
 aaacagtggt agtgttttgg tgatcatttg ctttgcaaca agccactccc caaagttagt 720
 ggcctaaaac aatttaataca cagttcatgt tctggctaca acaatacaca tccctctcat 780
 gtgcaaaata cactcactcc tccctcagag cctcgtagca ttaagggttc aggttcaaaag 840
 cttaagatct tatcctctga agtaggttta gggacaaaca agtcttctca ggtactctct 900
 ctggggagac agagacttgt gaactaaaag acaagttacc taccttccaa cacaactgac 960
 atgcaatggg gatataggaa aagataattt caataggcgc ttctgtgcaa aagcggggga 1020
 aatgagagtc actcagcagt cacggttcat attaatctaa aatctagcca ggcataatc 1080
 ccaagtcttc ctgatgtgag gacaagaatt attctctgat tagggctcac ttwttctctt 1140
 tgaggntggg tcgcctcagc ttttggattt gtctctgaa tcactcttcc ttgtctataa 1200
 aatgcatgta tatactcata catacataga gagaaagaga gagagagaga gagagagact 1260
 ctgtcacgca ggctggagtg caatgggtgt atctcagctc actgcaacct acaactcctg 1320
 ggttcaagca attctcctgt ctcagcctcc cgagcacctg tagtccctgc tactcaggag 1380
 gctgaggcag gagaattgct tgaatccgag aggcagaggt tgtcagttag cagagattac 1440
 accactgcac tccagcttgg gtgacagagc aaggcttcat ctcaaaaaaa gacaaaaaaa 1500
 aaaaaaaaaa actcgtag 1518

<210> 110
 <211> 921
 <212> DNA
 <213> Homo sapiens

<400> 110
 ggcacgagac gccgtgagcg ggacgaccag atgtaccgag agcgggctgcg caccttgctg 60
 gtcacgcggt ttgtcatgag cctcctgaat gctctcagca ccagcggagg cagcatttcc 120
 tggaaagact ttgtccacga gatgtggcc aagggcgagg tgcagcgcgt ccaggtgggtg 180
 cctgagagcg acgtggtgga agtctacctg caccctggag ccgtggtgtt tgggaggcct 240
 cggctagcct tgatgtaccg aatgcagttg caaatattga caagtttgaa gagaagcttc 300
 gagcagctga agatgagctg aatatcgagg ccaaggacag gatccagtt tcctacaagc 360
 gaacaggatt ctttgggaaa tgcctgttac tctgtgggga tgacggtagt gggcctggcc 420
 atctgtggtt atgttttccg tctggccggg atgactggag gcaccgccgg cgatggagct 480
 ccatgtcccg gtcctgtgct tggaaagcgt tgatggggag cgtcggcgct gaccacacgc 540
 gggagctgcg gaagcccagc ggttcacaca ggcctccctt caacgtagtc atcccctggt 600
 ggtggaagca agacgacggc ccttgacgtg cagccacaca cagaaaaggc tgctgtgaac 660
 attttatgct tcgacttttt tttctctcag agacagggtg tcgttctgtc gccaggctg 720
 gagtgcagtg ccaccatcat agctcactgc agcctccacc tcttaggctc aagcttccca 780
 agtagttggg actcaaggct tgagtaccca tgccaggctc tgttttttca gtctgtgaaa 840
 aataaagtca tcagcatgtg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
 aaaaaaaaaa aaaaaaaaaa a 921

<210> 111
 <211> 244
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (244)
 <223> Xaa equals stop translation

<400> 111
 Met Gly Thr Leu Pro Trp Leu Leu Ala Phe Phe Ile Leu Gly Leu Gln
 1 5 10 15
 Ala Trp Asp Thr Pro Thr Ile Val Ser Arg Lys Glu Trp Gly Ala Arg
 20 25 30
 Pro Leu Ala Cys Arg Ala Leu Leu Thr Leu Pro Val Ala Tyr Ile Ile
 35 40 45
 Thr Asp Gln Leu Pro Gly Met Gln Cys Gln Gln Gln Ser Val Cys Ser
 50 55 60
 Gln Met Leu Arg Gly Leu Gln Ser His Ser Val Tyr Thr Ile Gly Trp
 65 70 75 80
 Cys Asp Val Ala Tyr Asn Phe Leu Val Gly Asp Asp Gly Arg Val Tyr
 85 90 95
 Glu Gly Val Gly Trp Asn Ile Gln Gly Leu His Thr Gln Gly Tyr Asn
 100 105 110
 Asn Ile Ser Leu Gly Ile Ala Phe Phe Gly Asn Lys Ile Ser Ser Ser
 115 120 125
 Pro Ser Pro Ala Ala Leu Ser Ala Ala Glu Gly Leu Ile Ser Tyr Ala
 130 135 140
 Ile Gln Lys Gly His Leu Ser Pro Arg Tyr Ile Gln Pro Leu Leu Leu
 145 150 155 160
 Lys Glu Glu Thr Cys Leu Asp Pro Gln His Pro Val Met Pro Arg Lys
 165 170 175
 Val Cys Pro Asn Ile Ile Lys Arg Ser Ala Trp Glu Ala Arg Glu Thr
 180 185 190
 His Cys Pro Lys Met Asn Leu Pro Ala Lys Tyr Val Ile Ile Ile His
 195 200 205
 Thr Ala Gly Thr Ser Cys Thr Val Ser Thr Asp Cys Gln Thr Val Val
 210 215 220
 Arg Asn Ile Gln Ser Phe His Met Asp Thr Arg Asn Phe Cys Asp Ile
 225 230 235 240
 Gly Tyr Gln Xaa

<210> 112
 <211> 42
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

<400> 112
 Met Lys Arg Arg Glu Met Thr Gln Phe Leu Leu Ser Leu Val Ala Leu
 1 5 10 15
 Asn Cys Cys Ser Ile Ser Leu Gly Arg Leu Thr Tyr Pro Gly Gly Phe
 20 25 30
 His Leu Lys Leu Asp Pro Leu Glu Leu Xaa
 35 40

<210> 113
 <211> 527
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (466)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (527)
 <223> Xaa equals stop translation

<400> 113
 Met Ala Ala Leu Thr Ile Ala Thr Gly Thr Gly Asn Trp Phe Ser Ala
 1 5 10 15
 Leu Ala Leu Gly Val Thr Leu Leu Lys Cys Leu Leu Ile Pro Thr Tyr
 20 25 30
 His Ser Thr Asp Phe Glu Val His Arg Asn Trp Leu Ala Ile Thr His
 35 40 45
 Ser Leu Pro Ile Ser Gln Trp Tyr Tyr Glu Ala Thr Ser Glu Trp Thr
 50 55 60
 Leu Asp Tyr Pro Pro Phe Phe Ala Trp Phe Glu Tyr Ile Leu Ser His
 65 70 75 80
 Val Ala Lys Tyr Phe Asp Gln Glu Met Leu Asn Val His Asn Leu Asn
 85 90 95
 Tyr Ser Ser Ser Arg Thr Leu Leu Phe Gln Arg Phe Ser Val Ile Phe
 100 105 110
 Met Asp Val Leu Phe Val Tyr Ala Val Arg Glu Cys Cys Lys Cys Ile
 115 120 125
 Asp Gly Lys Lys Val Gly Lys Glu Leu Thr Glu Lys Pro Lys Phe Ile
 130 135 140
 Leu Ser Val Leu Leu Leu Trp Asn Phe Gly Leu Leu Ile Val Asp His
 145 150 155 160
 Ile His Phe Gln Tyr Asn Gly Phe Leu Phe Gly Leu Met Leu Leu Ser

165										170					175				
Ile	Ala	Arg	Leu	Phe	Gln	Lys	Arg	His	Met	Glu	Gly	Ala	Phe	Leu	Phe				
			180					185					190						
Ala	Val	Leu	Leu	His	Phe	Lys	His	Ile	Tyr	Leu	Tyr	Val	Ala	Pro	Ala				
		195					200					205							
Tyr	Gly	Val	Tyr	Leu	Leu	Arg	Ser	Tyr	Cys	Phe	Thr	Ala	Asn	Lys	Pro				
	210					215					220								
Asp	Gly	Ser	Ile	Arg	Trp	Lys	Ser	Phe	Ser	Phe	Val	Arg	Val	Ile	Ser				
225					230					235				240					
Leu	Gly	Leu	Val	Val	Phe	Leu	Val	Ser	Ala	Leu	Ser	Leu	Gly	Pro	Phe				
			245					250					255						
Leu	Ala	Leu	Asn	Gln	Leu	Pro	Gln	Val	Phe	Ser	Arg	Leu	Phe	Pro	Phe				
		260					265						270						
Lys	Arg	Gly	Leu	Cys	His	Ala	Tyr	Trp	Ala	Pro	Asn	Phe	Trp	Ala	Leu				
	275					280					285								
Tyr	Asn	Ala	Leu	Asp	Lys	Val	Leu	Ser	Val	Ile	Gly	Leu	Lys	Leu	Lys				
	290				295					300									
Phe	Leu	Asp	Pro	Asn	Asn	Ile	Pro	Lys	Ala	Ser	Met	Thr	Ser	Gly	Leu				
305				310					315					320					
Val	Gln	Gln	Phe	Gln	His	Thr	Val	Leu	Pro	Ser	Val	Thr	Pro	Leu	Ala				
			325				330						335						
Thr	Leu	Ile	Cys	Thr	Leu	Ile	Ala	Ile	Leu	Pro	Ser	Ile	Phe	Cys	Leu				
		340				345						350							
Trp	Phe	Lys	Pro	Gln	Gly	Pro	Arg	Gly	Phe	Leu	Arg	Cys	Leu	Thr	Leu				
	355				360						365								
Cys	Ala	Leu	Ser	Ser	Phe	Met	Phe	Gly	Trp	His	Val	His	Glu	Lys	Ala				
	370				375					380									
Ile	Leu	Leu	Ala	Ile	Leu	Pro	Met	Ser	Leu	Leu	Ser	Val	Gly	Lys	Ala				
385				390					395					400					
Gly	Asp	Ala	Ser	Ile	Phe	Leu	Ile	Leu	Thr	Thr	Thr	Gly	His	Tyr	Ser				
			405					410					415						
Leu	Phe	Pro	Leu	Leu	Phe	Thr	Ala	Pro	Glu	Leu	Pro	Ile	Lys	Ile	Leu				
		420					425					430							
Leu	Met	Leu	Leu	Phe	Thr	Ile	Tyr	Ser	Ile	Ser	Ser	Leu	Lys	Thr	Leu				
	435					440					445								
Phe	Arg	Lys	Glu	Lys	Pro	Leu	Phe	Asn	Trp	Met	Glu	Thr	Phe	Tyr	Leu				
	450				455					460									
Leu	Xaa	Leu	Gly	Pro	Leu	Glu	Val	Cys	Cys	Glu	Phe	Val	Phe	Pro	Phe				
465				470					475					480					
Thr	Ser	Trp	Lys	Val	Lys	Tyr	Pro	Phe	Ile	Pro	Leu	Leu	Leu	Thr	Ser				
			485				490						495						
Val	Tyr	Cys	Ala	Val	Gly	Ile	Thr	Tyr	Ala	Trp	Phe	Lys	Leu	Tyr	Val				
		500					505						510						

Ser Val Leu Ile Asp Ser Ala Ile Gly Lys Thr Lys Lys Gln Xaa
515 520 525

<210> 114
<211> 354
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (98)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (100)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (109)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (123)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (129)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (131)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (159)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (169)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (171)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (172)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (175)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (183)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (188)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (189)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (225)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (229)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (231)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 114
 Met Glu Asp Gly Val Leu Lys Glu Gly Phe Leu Val Lys Arg Gly His
 1 5 10 15
 Ile Val His Asn Trp Lys Ala Arg Trp Phe Ile Leu Arg Gln Asn Thr
 20 25 30
 Leu Val Tyr Tyr Lys Leu Glu Gly Gly Arg Arg Val Thr Pro Pro Lys
 35 40 45
 Gly Arg Ile Leu Leu Asp Gly Cys Thr Ile Thr Cys Pro Cys Leu Glu
 50 55 60
 Tyr Glu Asn Arg Pro Leu Leu Ile Lys Leu Lys Thr Gln Thr Ser Thr
 65 70 75 80
 Glu Tyr Phe Leu Glu Ala Cys Ser Arg Glu Glu Ala Gly Cys Leu Gly
 85 90 95
 Leu Xaa Arg Xaa Pro Gly Leu Phe Met Gln Gly Ser Xaa Gly Lys Val
 100 105 110
 Gln Gln Leu His Ser Leu Arg Asn Ser Phe Xaa Leu Pro Pro His Ile
 115 120 125
 Xaa Leu Xaa Arg Ile Val Asp Lys Met His Asp Ser Asn Thr Gly Ile
 130 135 140
 Arg Ser Ser Pro Asn Met Glu Gln Arg Ser Thr Tyr Lys Lys Xaa Phe
 145 150 155 160
 Leu Gly Ser Ser Leu Val Asp Trp Xaa Ile Xaa Xaa Ser Phe Xaa Gly
 165 170 175

Ser Arg Leu Glu Ala Val Xaa Leu Ala Ser Met Xaa Xaa Glu Glu Asn
 180 185 190
 Phe Leu Arg Ser Val Ala Val Arg Cys Met Gly Gly Ile Arg Ser Gly
 195 200 205
 Asp Leu Ala Glu Gln Phe Leu Asp Asp Ser Thr Ala Leu Tyr Thr Phe
 210 215 220
 Xaa Glu Ser Tyr Xaa Lys Xaa Ile Ser Pro Lys Glu Glu Ile Ser Leu
 225 230 235 240
 Ser Thr Val Glu Leu Ser Gly Thr Val Val Lys Gln Gly Tyr Leu Ala
 245 250 255
 Lys Gln Gly His Lys Arg Lys Asn Trp Lys Val Arg Arg Phe Val Leu
 260 265 270
 Arg Lys Asp Pro Ala Phe Leu His Tyr Tyr Asp Pro Ser Lys Glu Glu
 275 280 285
 Asn Arg Pro Val Gly Gly Phe Ser Leu Arg Gly Ser Leu Val Ser Ala
 290 295 300
 Leu Glu Asp Asn Gly Val Pro Thr Gly Val Lys Gly Asn Val Gln Gly
 305 310 315 320
 Asn Leu Phe Lys Val Ile Thr Lys Asp Asp Thr His Tyr Tyr Ile Gln
 325 330 335
 Ala Ser Ser Lys Ala Glu Arg Ala Glu Trp Ile Glu Ala Ile Lys Lys
 340 345 350
 Leu Thr

<210> 115
 <211> 64
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (64)
 <223> Xaa equals stop translation

<400> 115
 Met Trp Lys Arg Val Cys Val Cys Val Phe Leu Tyr Ile Ala Trp Val
 1 5 10 15
 Gln Leu Trp Met Cys Ala Lys Glu Cys Glu Cys Val Cys Val Cys Val
 20 25 30
 Lys Gly Ser Val Leu Glu Pro Thr Ser Val Cys Cys Glu Ser Gly Lys
 35 40 45
 Arg Val Gly Glu Gly Arg Glu Met Leu Thr Leu Val Gly Ala Gly Xaa
 50 55 60

```
<220>
<221> SITE
<222> (129)
<223> Xaa equals any of the naturally occurring L-amino acids.
```

```
<220>
<221> SITE
<222> (178)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>
<221> SITE
<222> (187)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>
<221> SITE
<222> (262)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>
<221> SITE
<222> (308)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>  
<221> SITE  
<222> (310)  
<223> Xaa equals stop translation
```

<400> 116															
Met	Phe	Thr	Ile	Lys	Leu	Leu	Leu	Phe	Ile	Val	Pro	Leu	Val	Ile	Ser
1				5					10					15	
Ser	Arg	Ile	Asp	Gln	Asp	Asn	Ser	Ser	Phe	Asp	Ser	Leu	Ser	Pro	Glu
			20					25					30		
Pro	Lys	Ser	Arg	Phe	Ala	Met	Leu	Asp	Asp	Val	Lys	Ile	Leu	Ala	Asn
		35					40					45			
Gly	Leu	Leu	Gln	Leu	Gly	His	Gly	Leu	Lys	Asp	Phe	Val	His	Lys	Thr
	50					55					60				
Lys	Gly	Gln	Ile	Asn	Asp	Ile	Phe	Gln	Lys	Leu	Asn	Ile	Phe	Asn	Gln
65					70					75					80
Ser	Phe	Tyr	Asp	Leu	Ser	Leu	Gln	Thr	Ser	Glu	Ile	Lys	Glu	Glu	Glu
				85					90					95	
Lys	Glu	Leu	Arg	Arg	Thr	Thr	Tyr	Lys	Leu	Gln	Val	Lys	Asn	Glu	Glu
			100					105					110		
Val	Lys	Asn	Met	Ser	Leu	Glu	Leu	Asn	Ser	Lys	Leu	Glu	Ser	Leu	Leu
		115					120					125			
Xaa	Glu	Lys	Ile	Leu	Leu	Gln	Gln	Lys	Val	Lys	Tyr	Leu	Glu	Glu	Gln
	130					135					140				

Leu Thr Asn Leu Ile Gln Asn Gln Pro Glu Thr Pro Glu His Pro Glu
 145 150 155 160
 Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser Ile Lys
 165 170 175
 Asp Xaa Leu Gln Thr Val Glu Asp Gln Tyr Xaa Gln Leu Asn Gln Gln
 180 185 190
 His Ser Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser Ile
 195 200 205
 Gln Glu Pro Thr Glu Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro Arg
 210 215 220
 Thr Thr Pro Phe Leu Gln Leu Asn Glu Ile Arg Asn Val Lys His Asp
 225 230 235 240
 Gly Ile Pro Ala Glu Cys Thr Thr Ile Tyr Asn Arg Gly Glu His Thr
 245 250 255
 Ser Gly Met Tyr Ala Xaa Arg Pro Ser Asn Ser Gln Val Phe His Val
 260 265 270
 Tyr Cys Asp Val Ile Ser Gly Ser Pro Trp Thr Leu Ile Gln His Arg
 275 280 285
 Ile Asp Gly Ser Gln Asn Phe Asn Glu Thr Trp Glu Asn Tyr Lys Tyr
 290 295 300
 Gly Phe Gly Xaa Ala Xaa
 305 310

<210> 117
 <211> 212
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (99)
 <223> Xaa equals any of the naturally occurring L-amino acids
 <220>
 <221> SITE
 <222> (212)
 <223> Xaa equals stop translation

<400> 117
 Met Ala Asn Ala Gly Leu Gln Leu Leu Gly Phe Ile Leu Ala Phe Leu
 1 5 10 15
 Gly Trp Ile Gly Ala Ile Val Ser Thr Ala Leu Pro Gln Trp Arg Ile
 20 25 30
 Tyr Ser Tyr Ala Gly Asp Asn Ile Val Thr Ala Gln Ala Met Tyr Glu
 35 40 45
 Gly Leu Trp Met Ser Cys Val Ser Gln Ser Thr Gly Gln Ile Gln Cys
 50 55 60
 Lys Val Phe Asp Ser Leu Leu Asn Leu Ser Ser Thr Leu Gln Ala Thr
 65 70 75 80

```
<210> 118
<211> 51
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation
```

<400> 118
Met Ala Pro Leu Trp Thr Leu Arg Pro Val Leu Val Trp Thr Thr Pro
1 5 10 15

Thr Ser Met Gly Glu Val Ser Pro Trp Leu Thr Ser Thr Val Met Ala
20 25 30

Lys Trp Thr Ser Ser Met Ala Thr Gly Met Ala Pro Thr Ala Ser Ile
35 40 45

Cys Arg Xaa
50

```
<210> 119
<211> 263
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (263)
<223> Xaa equals stop translation
<400> 119
```


Met Leu Phe Ser Ala Leu Leu Leu Glu Val Ile Trp Ile Leu Ala Ala
 1 5 10 15
 Asp Gly Gly Gln His Trp Thr Tyr Glu Gly Pro His Gly Gln Asp His
 20 25 30
 Trp Pro Ala Ser Tyr Pro Glu Cys Gly Asn Asn Ala Gln Ser Pro Ile
 35 40 45
 Asp Ile Gln Thr Asp Ser Val Thr Phe Asp Pro Asp Leu Pro Ala Leu
 50 55 60
 Gln Pro His Gly Tyr Asp Gln Pro Gly Thr Glu Pro Leu Asp Leu His
 65 70 75 80
 Asn Asn Gly His Thr Val Gln Leu Ser Leu Pro Ser Thr Leu Tyr Leu
 85 90 95
 Gly Gly Leu Pro Arg Lys Tyr Val Ala Ala Gln Leu His Leu His Trp
 100 105 110
 Gly Gln Lys Gly Ser Pro Gly Gly Ser Glu His Gln Ile Asn Ser Glu
 115 120 125
 Ala Thr Phe Ala Glu Leu His Ile Val His Tyr Asp Ser Asp Ser Tyr
 130 135 140
 Asp Ser Leu Ser Glu Ala Ala Glu Arg Pro Gln Gly Leu Ala Val Leu
 145 150 155 160
 Gly Ile Leu Ile Glu Leu Glu Lys Leu Gln Gly Thr Leu Phe Ser Thr
 165 170 175
 Glu Glu Glu Pro Ser Lys Leu Leu Val Gln Asn Tyr Arg Ala Leu Gln
 180 185 190
 Pro Leu Asn Gln Arg Met Val Phe Ala Ser Phe Ile Gln Ala Gly Ser
 195 200 205
 Ser Tyr Thr Thr Gly Glu Met Leu Ser Leu Gly Val Gly Ile Leu Val
 210 215 220
 Gly Cys Leu Cys Leu Leu Leu Ala Val Tyr Phe Ile Ala Arg Lys Ile
 225 230 235 240
 Arg Lys Lys Arg Leu Glu Asn Arg Lys Ser Val Val Phe Thr Ser Ala
 245 250 255
 Gln Ala Thr Thr Glu Ala Xaa
 260

<210> 120

<211> 270

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (27)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 120

Met His Tyr Tyr Arg Tyr Ser Asn Ala Lys Val Ser Cys Trp Tyr Lys

72

1 5 10 15
 Tyr Leu Leu Phe Ser Tyr Asn Ile Ile Phe Xaa Leu Ala Gly Val Val
 20 25 30
 Phe Leu Gly Val Gly Leu Trp Ala Trp Ser Glu Lys Gly Val Leu Ser
 35 40 45
 Asp Leu Thr Lys Val Thr Arg Met His Gly Ile Asp Pro Val Val Leu
 50 55 60
 Val Leu Met Val Gly Val Val Met Phe Thr Leu Gly Phe Ala Gly Cys
 65 70 75 80
 Val Gly Ala Leu Arg Glu Asn Ile Cys Leu Leu Asn Phe Phe Cys Gly
 85 90 95
 Thr Ile Val Leu Ile Phe Phe Leu Glu Leu Ala Val Ala Val Leu Ala
 100 105 110
 Phe Leu Phe Gln Asp Trp Val Arg Asp Arg Phe Arg Glu Phe Phe Glu
 115 120 125
 Ser Asn Ile Lys Ser Tyr Arg Asp Asp Ile Asp Leu Gln Asn Leu Ile
 130 135 140
 Asp Ser Leu Gln Lys Ala Asn Gln Cys Cys Gly Ala Tyr Gly Pro Glu
 145 150 155 160
 Asp Trp Asp Leu Asn Val Tyr Phe Asn Cys Ser Gly Ala Ser Tyr Ser
 165 170 175
 Arg Glu Lys Cys Gly Val Pro Phe Ser Cys Cys Val Pro Asp Pro Ala
 180 185 190
 Gln Lys Val Val Asn Thr Gln Cys Gly Tyr Asp Val Arg Ile Gln Leu
 195 200 205
 Lys Ser Lys Trp Asp Glu Ser Ile Phe Thr Lys Gly Cys Ile Gln Ala
 210 215 220
 Leu Glu Ser Trp Leu Pro Arg Asn Ile Tyr Ile Val Ala Gly Val Phe
 225 230 235 240
 Ile Ala Ile Ser Leu Leu Gln Ile Phe Gly Ile Phe Leu Ala Arg Thr
 245 250 255
 Leu Ile Ser Asp Ile Glu Ala Val Lys Ala Gly His His Phe
 260 265 270

<210> 121

<211> 92

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (92)

<223> Xaa equals stop translation

<400> 121

Met Leu Arg Cys Gly Gly Arg Gly Leu Leu Leu Gly Leu Ala Val Ala
 1 5 10 15

Ala Ala Ala Val Met Ala Ala Arg Leu Met Gly Trp Trp Gly Pro Arg
 20 25 30

Ala Gly Phe Arg Leu Phe Ile Pro Glu Glu Leu Ser Arg Tyr Arg Gly
 35 40 45

Gly Pro Gly Asp Pro Gly Leu Tyr Leu Ala Leu Leu Gly Arg Val Tyr
 50 55 60

Asp Val Ser Ser Gly Arg Ser Thr Thr Ser Leu Gly Pro Thr Ile Ala
 65 70 75 80

Ala Ser Gln Ala Glu Thr His Pro Glu Leu Ser Xaa
 85 90

<210> 122
 <211> 223
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (120)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (223)
 <223> Xaa equals stop translation

<400> 122
 Met Leu Trp Leu Leu Phe Phe Leu Val Thr Ala Ile His Ala Glu Leu
 1 5 10 15

Cys Gln Pro Gly Ala Glu Asn Ala Phe Lys Val Arg Leu Ser Ile Arg
 20 25 30

Thr Ala Leu Gly Asp Lys Ala Tyr Ala Trp Asp Thr Asn Glu Glu Tyr
 35 40 45

Leu Phe Lys Ala Met Val Ala Phe Ser Met Arg Lys Val Pro Asn Arg
 50 55 60

Glu Ala Thr Glu Ile Ser His Val Leu Leu Cys Asn Val Thr Gln Arg
 65 70 75 80

Val Ser Phe Trp Phe Val Val Thr Asp Pro Ser Lys Asn His Thr Leu
 85 90 95

Pro Ala Val Glu Val Gln Ser Ala Ile Arg Met Asn Lys Asn Arg Ile
 100 105 110

Asn Asn Ala Phe Phe Leu Asn Xaa Gln Thr Leu Glu Phe Leu Lys Ile
 115 120 125

Pro Ser Thr Leu Ala Pro Pro Met Asp Pro Ser Val Pro Ile Trp Ile
 130 135 140

Ile Ile Phe Gly Val Ile Phe Cys Ile Ile Ile Val Ala Ile Ala Leu
 145 150 155 160

Leu Ile Leu Ser Gly Ile Trp Gln Arg Arg Arg Lys Asn Lys Glu Pro

Asn Asp Ala Phe Met Thr Glu Asp Glu Arg Leu Thr Pro Leu Xaa
210 215 220

<210> 124

<211> 47
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (47)
 <223> Xaa equals stop translation

<400> 124
 Met Gly Arg Gln Ala Leu Leu Leu Leu Ala Leu Cys Ala Thr Gly Ala
 1 5 10 15
 Gln Gly Leu Tyr Phe His Ile Gly Glu Thr Glu Lys Arg Cys Phe Ile
 20 25 30
 Glu Glu Ile Pro Asp Glu Thr Met Val Ile Gly Gln Ala Gly Xaa
 35 40 45

<210> 125
 <211> 306
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (11)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (306)
 <223> Xaa equals stop translation

<400> 125
 Met Ala Leu Cys Ala Leu Thr Arg Ala Leu Xaa Ser Leu Asn Leu Ala
 1 5 10 15
 Pro Pro Thr Val Ala Ala Pro Ala Pro Ser Leu Phe Pro Ala Ala Gln
 20 25 30
 Met Met Asn Asn Gly Leu Leu Gln Gln Pro Ser Ala Leu Met Leu Leu
 35 40 45
 Pro Cys Arg Pro Val Leu Thr Ser Val Ala Leu Asn Ala Asn Phe Val
 50 55 60
 Ser Trp Lys Ser Arg Thr Lys Tyr Thr Ile Thr Pro Val Lys Met Arg
 65 70 75 80
 Lys Ser Gly Gly Arg Asp His Thr Gly Arg Ile Arg Val His Gly Ile
 85 90 95
 Gly Gly Gly His Lys Gln Arg Tyr Arg Met Ile Asp Phe Leu Arg Phe
 100 105 110
 Arg Pro Glu Glu Thr Lys Ser Gly Pro Phe Glu Glu Lys Val Ile Gln
 115 120 125
 Val Arg Tyr Asp Pro Cys Arg Ser Ala Asp Ile Ala Leu Val Ala Gly
 130 135 140
 Gly Ser Arg Lys Arg Trp Ile Ile Ala Thr Glu Asn Met Gln Ala Gly

```
<210> 126
<211> 82
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (82)
<223> Xaa equals stop translation
```

```

<400> 126
Met Asn Gln Leu Met Phe Gln Asp Leu Leu Cys Cys Leu Cys Leu Phe
  1                      5                      10                      15

Val Ile Gly Leu Ile Ser Leu Leu Arg Lys Thr Tyr Ser Cys Val Asn
      20                      25                      30

Leu Cys Lys Val Met Leu Pro Val Lys Lys Tyr Ser Thr Val Ser Thr
    35                      40                      45

Val Leu Cys Arg Asn Met Lys Leu Asn Gly Lys Asn Val Leu Met Phe
    50                      55                      60

Val Val Met Leu Leu Gly Gln Trp Met Gly Lys Leu Pro Lys Leu Ser
    65                      70                      75                      80

Pro Xaa

```

<210> 127

<211> 243
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (88)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (139)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (243)
 <223> Xaa equals stop translation

<400> 127
 Met Glu Gln Ala Arg Lys Ser Ser Thr Val Ser Leu Leu Ile Thr Val
 1 5 10 15
 Leu Phe Ala Val Ala Phe Ser Val Leu Leu Ser Cys Lys Asp His
 20 25 30
 Val Gly Tyr Ile Phe Thr Thr Asp Arg Asp Ile Ile Asn Leu Val Ala
 35 40 45
 Gln Val Val Pro Ile Tyr Ala Val Ser His Leu Phe Glu Ala Leu Ala
 50 55 60
 Cys Thr Ser Gly Gly Val Leu Arg Gly Ser Gly Asn Gln Lys Val Gly
 65 70 75 80
 Ala Ile Val Asn Thr Ile Gly Xaa Tyr Val Val Gly Leu Pro Ile Gly
 85 90 95
 Ile Ala Leu Met Phe Ala Thr Thr Leu Gly Val Met Gly Leu Trp Ser
 100 105 110
 Gly Ile Ile Ile Cys Thr Val Phe Gln Ala Val Cys Phe Leu Gly Phe
 115 120 125
 Ile Ile Gln Leu Asn Trp Lys Lys Ala Cys Xaa Gln Ala Gln Val His
 130 135 140
 Ala Asn Leu Lys Val Asn Asn Val Pro Arg Ser Gly Asn Ser Ala Leu
 145 150 155 160
 Pro Gln Asp Pro Leu His Pro Gly Cys Pro Glu Asn Leu Glu Gly Ile
 165 170 175
 Leu Thr Asn Asp Val Gly Lys Thr Gly Glu Pro Gln Ser Asp Gln Gln
 180 185 190
 Met Arg Gln Glu Glu Pro Leu Pro Glu His Pro Gln Asp Gly Ala Lys
 195 200 205
 Leu Ser Arg Lys Gln Leu Val Leu Arg Arg Gly Leu Leu Leu Gly
 210 215 220
 Val Phe Leu Ile Leu Leu Val Gly Ile Leu Val Arg Phe Tyr Val Arg
 225 230 235 240

Ile Gln Xaa

<210> 128

<211> 285

<212> PRT

<213> Homo sapiens

<400> 128

Met Val Val Ala Gly Val Val Val Leu Ile Leu Ala Leu Val Leu Ala
 1 5 10 15

Trp Leu Ser Thr Tyr Val Ala Asp Ser Gly Ser Asn Gln Leu Leu Gly
 20 25 30

Ala Ile Val Ser Ala Gly Asp Thr Ser Val Leu His Leu Gly His Val
 35 40 45

Asp His Leu Val Ala Gly Gln Gly Asn Pro Glu Pro Thr Glu Leu Pro
 50 55 60

His Pro Ser Glu Gly Asn Asp Glu Lys Ala Glu Glu Ala Gly Glu Gly
 65 70 75 80

Arg Gly Asp Ser Thr Gly Glu Ala Gly Ala Gly Gly Gly Val Glu Pro
 85 90 95

Ser Leu Glu His Leu Leu Asp Ile Gln Gly Leu Pro Lys Arg Gln Ala
 100 105 110

Gly Ala Gly Ser Ser Ser Pro Glu Ala Pro Leu Arg Ser Glu Asp Ser
 115 120 125

Thr Cys Leu Pro Pro Ser Pro Gly Leu Ile Thr Val Arg Leu Lys Phe
 130 135 140

Leu Asn Asp Thr Glu Glu Leu Ala Val Ala Arg Pro Glu Asp Thr Val
 145 150 155 160

Gly Ala Leu Lys Ser Lys Tyr Phe Pro Gly Gln Glu Ser Gln Met Lys
 165 170 175

Leu Ile Tyr Gln Gly Arg Leu Leu Gln Asp Pro Ala Arg Thr Leu Arg
 180 185 190

Ser Leu Asn Ile Thr Asp Asn Cys Val Ile His Cys His Arg Ser Pro
 195 200 205

Pro Gly Ser Ala Val Pro Gly Pro Ser Ala Ser Leu Ala Pro Ser Ala
 210 215 220

Thr Glu Pro Pro Ser Leu Gly Val Asn Val Gly Ser Leu Met Val Pro
 225 230 235 240

Val Phe Val Val Leu Leu Gly Val Val Trp Tyr Phe Arg Ile Asn Tyr
 245 250 255

Arg Gln Phe Phe Thr Ala Pro Ala Thr Val Ser Leu Val Gly Val Thr
 260 265 270

Val Phe Phe Ser Phe Leu Val Phe Gly Met Tyr Gly Arg
 275 280 285

<210> 129
 <211> 158
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (114)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (119)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (120)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (121)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (158)
 <223> Xaa equals stop translation

<400> 129
 Met Asp Ala Met Ile Leu Leu Asn Val Leu Ala Leu Thr Arg Leu Ala
 1 5 10 15
 Lys Ala Ala Ala Thr Asn Phe Val Ala Gln Gly Arg Gly Thr Ile Ile
 20 25 30
 Asn Ile Gly Ser Ile Val Ala Leu Ala Pro Lys Val Leu Asn Gly Val
 35 40 45
 Tyr Gly Gly Thr Lys Ala Phe Val Gln Ala Phe Ser Glu Ser Leu Gln
 50 55 60
 His Glu Leu Ser Asp Lys Gly Val Val Val Gln Val Val Leu Pro Gly
 65 70 75 80
 Ala Thr Ala Thr Glu Phe Trp Asp Ile Ala Gly Leu Pro Val Lys Gln
 85 90 95
 Pro Ala Gly Ser His Gly Asp Asp His Arg Lys Pro Gly Gly Arg Arg
 100 105 110
 Pro Xaa Arg Pro Cys Pro Xaa Xaa Xaa Val Thr Ile Pro Ser Leu Pro
 115 120 125
 Asp Ser Ala Asp Trp Asp Thr Thr Asn Ala Arg Gly Trp Pro Trp Val
 130 135 140
 Arg Thr Cys Arg Thr Val Asn Pro Pro Leu Val Met Gly Xaa
 145 150 155

<210> 130
 <211> 309
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (87)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (185)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (309)
 <223> Xaa equals stop translation

<400> 130
 Met Pro Val Pro Trp Phe Leu Leu Ser Leu Ala Leu Gly Arg Ser Pro
 1 5 10 15
 Val Val Leu Ser Leu Glu Arg Leu Val Gly Pro Gln Asp Ala Thr His
 20 25 30
 Cys Ser Pro Gly Leu Ser Cys Arg Leu Trp Asp Ser Asp Ile Leu Cys
 35 40 45
 Leu Pro Gly Asp Ile Val Pro Ala Pro Gly Pro Val Leu Ala Pro Thr
 50 55 60
 His Leu Gln Thr Glu Leu Val Leu Arg Cys Gln Lys Glu Thr Asp Cys
 65 70 75 80
 Asp Leu Cys Leu Arg Val Xaa Val His Leu Ala Val His Gly His Trp
 85 90 95
 Glu Glu Pro Glu Asp Glu Glu Lys Phe Gly Gly Ala Ala Asp Leu Gly
 100 105 110
 Val Glu Glu Pro Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu Ser
 115 120 125
 Phe Gln Ala Tyr Pro Thr Ala Arg Cys Val Leu Leu Glu Val Gln Val
 130 135 140
 Pro Ala Ala Leu Val Gln Phe Gly Gln Ser Val Gly Ser Val Val Tyr
 145 150 155 160
 Asp Cys Phe Glu Ala Ala Leu Gly Ser Glu Val Arg Ile Trp Ser Tyr
 165 170 175
 Thr Gln Pro Arg Tyr Glu Lys Glu Xaa Asn His Thr Gln Gln Leu Pro
 180 185 190
 Asp Cys Arg Gly Leu Glu Val Trp Asn Ser Ile Pro Ser Cys Trp Ala
 195 200 205
 Leu Pro Trp Leu Asn Val Ser Ala Asp Gly Asp Asn Val His Leu Val
 210 215 220
 Leu Asn Val Ser Glu Glu Gln His Phe Gly Leu Ser Leu Tyr Trp Asn

81

225 230 235 240
 Gln Val Gln Gly Pro Pro Lys Pro Arg Trp His Lys Asn Leu Thr Gly
 245 250 255
 Pro Gln Ile Ile Thr Leu Asn His Thr Asp Leu Val Pro Cys Leu Cys
 260 265 270
 Ile Gln Val Trp Pro Leu Glu Pro Asp Ser Val Arg Arg Thr Ser Ala
 275 280 285
 Pro Ser Gly Arg Thr Pro Ala His Thr Arg Thr Ser Gly Lys Pro Pro
 290 295 300
 Asp Cys Asp Cys Xaa
 305

<210> 131
 <211> 509
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (509)
 <223> Xaa equals stop translation

<400> 131
 Met Asp Pro Lys Leu Gly Arg Met Ala Ala Ser Leu Leu Ala Val Leu
 1 5 10 15
 Leu Leu Leu Leu Leu Glu Arg Gly Met Phe Ser Ser Pro Ser Pro Pro
 20 25 30
 Pro Ala Leu Leu Glu Lys Val Phe Gln Tyr Ile Asp Leu His Gln Asp
 35 40 45
 Glu Phe Val Gln Thr Leu Lys Glu Trp Val Ala Ile Glu Ser Asp Ser
 50 55 60
 Val Gln Pro Val Pro Arg Phe Arg Gln Glu Leu Phe Arg Met Met Ala
 65 70 75 80
 Val Ala Ala Asp Thr Leu Gln Arg Leu Gly Ala Arg Val Ala Ser Val
 85 90 95
 Asp Met Gly Pro Gln Gln Leu Pro Asp Gly Gln Ser Leu Pro Ile Pro
 100 105 110
 Pro Val Ile Leu Ala Glu Leu Gly Ser Asp Pro Thr Lys Gly Thr Val
 115 120 125
 Cys Phe Tyr Gly His Leu Asp Val Gln Pro Ala Asp Arg Gly Asp Gly
 130 135 140
 Trp Leu Thr Asp Pro Tyr Val Leu Thr Glu Val Asp Gly Lys Leu Tyr
 145 150 155 160
 Gly Arg Gly Ala Thr Asp Asn Lys Gly Pro Val Leu Ala Trp Ile Asn
 165 170 175
 Ala Val Ser Ala Phe Arg Ala Leu Glu Gln Asp Leu Pro Val Asn Ile
 180 185 190

Lys Phe Ile Ile Glu Gly Met Glu Glu Ala Gly Ser Val Ala Leu Glu
 195 200 205
 Glu Leu Val Glu Lys Glu Lys Asp Arg Phe Phe Ser Gly Val Asp Tyr
 210 215 220
 Ile Val Ile Ser Asp Asn Leu Trp Ile Ser Gln Arg Lys Pro Ala Ile
 225 230 235 240
 Thr Tyr Gly Thr Arg Gly Asn Ser Tyr Phe Met Val Glu Val Lys Cys
 245 250 255
 Arg Asp Gln Asp Phe His Ser Gly Thr Phe Gly Gly Ile Leu His Glu
 260 265 270
 Pro Met Ala Asp Leu Val Ala Leu Leu Gly Ser Leu Val Asp Ser Ser
 275 280 285
 Gly His Ile Leu Val Pro Gly Ile Tyr Asp Glu Val Val Pro Leu Thr
 290 295 300
 Glu Glu Glu Ile Asn Thr Tyr Lys Ala Ile His Leu Asp Leu Glu Glu
 305 310 315 320
 Tyr Arg Asn Ser Ser Arg Val Glu Lys Phe Leu Phe Asp Thr Lys Glu
 325 330 335
 Glu Ile Leu Met His Leu Trp Arg Tyr Pro Ser Leu Ser Ile His Gly
 340 345 350
 Ile Glu Gly Ala Phe Asp Glu Pro Gly Thr Lys Thr Val Ile Pro Gly
 355 360 365
 Arg Val Ile Gly Lys Phe Ser Ile Arg Leu Val Pro His Met Asn Val
 370 375 380
 Ser Ala Val Glu Lys Gln Val Thr Arg His Leu Glu Asp Val Phe Ser
 385 390 395 400
 Lys Arg Asn Ser Ser Asn Lys Met Val Val Ser Met Thr Leu Gly Leu
 405 410 415
 His Pro Trp Ile Ala Asn Ile Asp Asp Thr Gln Tyr Leu Ala Ala Lys
 420 425 430
 Arg Ala Ile Arg Thr Val Phe Gly Thr Glu Pro Asp Met Ile Arg Asp
 435 440 445
 Gly Ser Thr Ile Pro Ile Ala Lys Met Phe Gln Glu Ile Val His Lys
 450 455 460
 Ser Val Val Leu Ile Pro Leu Gly Ala Val Asp Asp Gly Glu His Ser
 465 470 475 480
 Gln Asn Glu Lys Ile Asn Arg Trp Asn Tyr Ile Glu Gly Thr Lys Leu
 485 490 495
 Phe Ala Ala Phe Phe Leu Glu Met Ala Gln Leu His Xaa
 500 505

<210> 132

<211> 507

<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (65)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (112)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (423)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (425)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (507)
<223> Xaa equals stop translation

<400> 132
Met Gly Met Arg Arg His Ser Leu Met Leu Leu Pro Trp Trp Leu Gly
1 5 10 15
Ala Ala Gly Arg Lys Glu Cys His Arg Glu Gln Leu Val Ala Ala Val
20 25 30
Glu Val Thr Glu Gln Glu Thr Lys Val Pro Lys Lys Thr Val Ile Ile
35 40 45
Glu Glu Thr Ile Thr Thr Val Val Lys Ser Pro Arg Gly Gln Arg Arg
50 55 60
Xaa Pro Ser Lys Ser Pro Ser Arg Ser Pro Ser Arg Cys Ser Ala Ser
65 70 75 80
Pro Leu Arg Pro Gly Leu Leu Ala Pro Asp Leu Leu Tyr Leu Pro Gly
85 90 95
Ala Gly Gln Pro Arg Arg Pro Glu Ala Glu Pro Gly Gln Lys Pro Xaa
100 105 110
Val Pro Thr Leu Tyr Val Thr Glu Ala Glu Ala His Ser Pro Ala Leu
115 120 125
Pro Gly Leu Ser Gly Pro Gln Pro Lys Trp Val Glu Val Glu Glu Thr
130 135 140
Ile Glu Val Arg Val Lys Lys Met Gly Pro Gln Gly Val Ser Pro Thr
145 150 155 160
Thr Glu Val Pro Arg Ser Ser Ser Gly His Leu Phe Thr Leu Pro Gly
165 170 175
Ala Thr Pro Gly Gly Asp Pro Asn Ser Asn Asn Ser Asn Asn Lys Leu
180 185 190

Leu Ala Gln Glu Ala Trp Ala Gln Gly Thr Ala Met Val Gly Val Arg
 195 200 205
 Glu Pro Leu Val Phe Arg Val Asp Ala Arg Gly Ser Val Asp Trp Ala
 210 215 220
 Ala Ser Gly Met Gly Ser Leu Glu Glu Gly Thr Met Glu Glu Ala
 225 230 235 240
 Gly Glu Glu Glu Gly Glu Asp Gly Asp Ala Phe Val Thr Glu Glu Ser
 245 250 255
 Gln Asp Thr His Ser Leu Gly Asp Arg Asp Pro Lys Ile Leu Thr His
 260 265 270
 Asn Gly Arg Met Leu Thr Leu Ala Asp Leu Glu Asp Tyr Val Pro Gly
 275 280 285
 Glu Gly Glu Thr Phe His Cys Gly Gly Pro Gly Pro Gly Ala Pro Asp
 290 295 300
 Asp Pro Pro Cys Glu Val Ser Val Ile Gln Arg Glu Ile Gly Glu Pro
 305 310 315 320
 Thr Val Gly Ser Leu Cys Cys Ser Ala Trp Gly Met His Trp Val Pro
 325 330 335
 Glu Ala Leu Ser Ala Ser Leu Gly Leu Ser Pro Val Gly Arg His His
 340 345 350
 Arg Asp Pro Arg Ser Val Ala Leu Arg Ala Pro Pro Ser Ser Cys Gly
 355 360 365
 Arg Pro Arg Leu Gly Leu Trp Ala Val Leu Pro Gly Arg Ser Leu Ser
 370 375 380
 Ala Pro Ala Ser Gly Val Leu Arg Thr Val Ala Arg Ala Ala Ser Pro
 385 390 395 400
 Gln Ser Phe Pro Pro Arg Pro Ser Thr Ser Gly Gln Trp Gly Arg Arg
 405 410 415
 Ser Pro Phe Thr Ser Val Xaa Gly Xaa Gly Pro Ser Tyr Leu Thr Gln
 420 425 430
 Leu Gln Pro Gly Gly Leu Gly Gly Ala Cys Asn Val Gly Met Thr Gly
 435 440 445
 Ser Lys Thr Ser Ala Leu Gly Cys Phe Leu Ser Ala Trp Gln Glu Pro
 450 455 460
 Gln Asp Cys Gly Arg Arg Met Trp Pro Trp Ala Phe Val Leu Phe Pro
 465 470 475 480
 His Gly Pro Gly Pro Ser Leu Leu Ala Pro Ala Thr Ala Ala Arg Pro
 485 490 495
 Asp Met Ala Leu Pro Leu Leu Gln Ser Trp Xaa
 500 505

<210> 133

<211> 49

<212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (49)
 <223> Xaa equals stop translation

 <400> 133
 Met Arg Leu Leu Leu Leu Leu Val Ala Ala Ser Ala Met Val Arg
 1 5 10 15

 Ser Glu Ala Ser Ala Asn Leu Gly Gly Val Pro Ser Lys Arg Leu Lys
 20 25 30

 Met Gln Tyr Ala Thr Gly Pro Leu Leu Lys Phe Gln Ile Cys Val Ser
 35 40 45

 Xaa

 <210> 134
 <211> 131
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (64)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (65)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (131)
 <223> Xaa equals stop translation

 <400> 134
 Met Leu Met Pro Val His Phe Leu Leu Leu Leu Leu Leu Leu Gly
 1 5 10 15

 Gly Pro Arg Thr Gly Leu Pro His Lys Phe Tyr Lys Ala Lys Pro Ile
 20 25 30

 Phe Ser Cys Leu Asn Thr Ala Leu Ser Glu Ala Glu Lys Gly Gln Trp
 35 40 45

 Glu Asp Ala Ser Leu Leu Ser Lys Arg Ser Phe His Tyr Leu Arg Xaa
 50 55 60

 Xaa Thr Pro Leu Arg Glu Arg Arg Arg Arg Ala Lys Arg Lys Arg Leu
 65 70 75 80

 Ser Pro Ser Leu Gly Pro Gly Val Glu Pro Glu Ala Pro Gly Thr Asp
 85 90 95

 Thr Cys Pro Lys His Ser Pro Gly Glu Ser His Ala Arg Thr Arg Pro
 100 105 110

Arg Val Pro Thr Ala Pro Ser Ser Pro Cys Pro Ser Thr Ser Pro Pro
 115 120 125

Thr Ser Xaa
 130

<210> 135
 <211> 44
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (25)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (29)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (44)
 <223> Xaa equals stop translation

<400> 135
 Met Ala Phe Leu Gln Ser Ala Ser Tyr Val Met Val Ile Leu Cys Ala
 1 5 10 15

Cys Val Ile Ile Ile Gly Ile Leu Xaa Tyr Ala Phe Xaa Phe Glu Thr
 20 25 30

Leu Ser Pro Lys Lys Arg Arg Asp Ile Glu Ile Xaa
 35 40

<210> 136
 <211> 92
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (92)
 <223> Xaa equals stop translation

<400> 136
 Met Gln Leu Ile Glu Ser Arg Phe His Phe Arg Cys Val Trp Ile Leu
 1 5 10 15

His Leu Leu Ala Leu Phe Ser Thr Trp Pro Pro Lys Asp Pro Glu Gly
 20 25 30

Ser Pro Pro Ser Ala Thr Ser Ser Pro Leu Thr Pro His Leu Ser Leu
 35 40 45

Thr Leu Pro Phe Lys Gln Ala Pro Val Ser Asn Val Ser Ser Ala Ile
 50 55 60

His Val Met Leu Asp Lys Ser Val Ser Leu Ser Glu Ile Gln Phe Ser
 65 70 75 80

His Met Pro Asn Gly Lys Arg Ala Ser Thr Leu Xaa
85 90

<210> 137
<211> 267
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (267)
<223> Xaa equals stop translation

<400> 137
Met Glu Leu Leu Thr Ala Leu Leu Arg Leu Phe Leu Ser Arg Pro Ala
1 5 10 15

Glu Cys Gln Asp Met Leu Gly Arg Leu Leu Tyr Tyr Cys Ile Glu Glu
20 25 30

Glu Lys Asp Met Ala Val Arg Asp Arg Gly Leu Phe Tyr Tyr Arg Leu
35 40 45

Leu Leu Val Gly Ile Asp Glu Val Lys Arg Ile Leu Cys Ser Pro Lys
50 55 60

Ser Asp Pro Thr Leu Gly Leu Leu Glu Asp Pro Ala Glu Arg Pro Val
65 70 75 80

Asn Ser Trp Ala Ser Asp Phe Asn Thr Leu Val Pro Val Tyr Gly Lys
85 90 95

Ala His Trp Ala Thr Ile Ser Lys Cys Gln Gly Ala Glu Arg Cys Asp
100 105 110

Pro Glu Leu Pro Lys Thr Ser Ser Phe Ala Ala Ser Gly Pro Leu Ile
115 120 125

Pro Glu Glu Asn Lys Glu Arg Val Gln Glu Leu Pro Asp Ser Gly Ala
130 135 140

Leu Met Leu Val Pro Asn Arg Gln Leu Thr Ala Asp Tyr Phe Glu Lys
145 150 155 160

Thr Trp Leu Ser Leu Lys Val Ala His Gln Gln Val Leu Pro Trp Arg
165 170 175

Gly Glu Phe His Pro Asp Thr Leu Gln Met Ala Leu Gln Val Val Asn
180 185 190

Ile Gln Thr Ile Ala Met Ser Arg Ala Gly Ser Arg Pro Trp Lys Ala
195 200 205

Tyr Leu Ser Ala Gln Asp Asp Thr Gly Cys Leu Phe Leu Thr Glu Leu
210 215 220

Leu Leu Glu Pro Gly Asn Ser Glu Met Gln Ile Ser Val Lys Gln Asn
225 230 235 240

Glu Ala Arg Thr Glu Thr Leu Asn Ser Phe Ile Ser Val Leu Glu Thr
245 250 255

Val Ile Gly Thr Ile Glu Glu Ile Lys Ser Xaa

260

265

<210> 138
 <211> 434
 <212> PRT
 <213> Homo sapiens

<400> 138

Met Ala Pro Glu Gly Leu Val Pro Ala Val Leu Trp Gly Leu Ser Leu
 1 5 10 15

Phe Leu Asn Leu Pro Gly Pro Ile Trp Leu Gln Pro Ser Pro Pro Pro
 20 25 30

Gln Ser Ser Pro Pro Pro Gln Pro His Pro Cys His Thr Cys Arg Gly
 35 40 45

Leu Val Asp Ser Phe Asn Lys Gly Leu Glu Arg Thr Ile Arg Asp Asn
 50 55 60

Phe Gly Gly Gly Asn Thr Ala Trp Glu Glu Glu Asn Leu Ser Lys Tyr
 65 70 75 80

Lys Asp Ser Glu Thr Arg Leu Val Glu Val Leu Glu Gly Val Cys Ser
 85 90 95

Lys Ser Asp Phe Glu Cys His Arg Leu Leu Glu Leu Ser Glu Glu Leu
 100 105 110

Val Glu Ser Trp Trp Phe His Lys Gln Gln Glu Ala Pro Asp Leu Phe
 115 120 125

Gln Trp Leu Cys Ser Asp Ser Leu Lys Leu Cys Cys Pro Ala Gly Thr
 130 135 140

Phe Gly Pro Ser Cys Leu Pro Cys Pro Gly Gly Thr Glu Arg Pro Cys
 145 150 155 160

Gly Gly Tyr Gly Gln Cys Glu Gly Glu Gly Thr Arg Gly Gly Ser Gly
 165 170 175

His Cys Asp Cys Gln Ala Gly Tyr Gly Gly Glu Ala Cys Gly Gln Cys
 180 185 190

Gly Leu Gly Tyr Phe Glu Ala Glu Arg Asn Ala Ser His Leu Val Cys
 195 200 205

Ser Ala Cys Phe Gly Pro Cys Ala Arg Cys Ser Gly Pro Glu Glu Ser
 210 215 220

Asn Cys Leu Gln Cys Lys Lys Gly Trp Ala Leu His His Leu Lys Cys
 225 230 235 240

Val Asp Ile Asp Glu Cys Gly Thr Glu Gly Ala Asn Cys Gly Ala Asp
 245 250 255

Gln Phe Cys Val Asn Thr Glu Gly Ser Tyr Glu Cys Arg Asp Cys Ala
 260 265 270

Lys Ala Cys Leu Gly Cys Met Gly Ala Gly Pro Gly Arg Cys Lys Lys
 275 280 285

Cys Ser Pro Gly Tyr Gln Gln Val Gly Ser Lys Cys Leu Asp Val Asp

290 295 300
 Glu Cys Glu Thr Glu Val Cys Pro Gly Glu Asn Lys Gln Cys Glu Asn
 305 310 315 320
 Thr Glu Gly Gly Tyr Arg Cys Ile Cys Ala Glu Gly Tyr Lys Gln Met
 325 330 335
 Glu Gly Ile Cys Val Lys Glu Gln Ile Pro Gly Ala Phe Pro Ile Leu
 340 345 350
 Thr Asp Leu Thr Pro Glu Thr Thr Arg Arg Trp Lys Leu Gly Ser His
 355 360 365
 Pro His Ser Thr Tyr Val Lys Met Lys Met Gln Arg Asp Glu Ala Thr
 370 375 380
 Phe Pro Gly Leu Tyr Gly Lys Gln Val Ala Lys Leu Gly Ser Gln Ser
 385 390 395 400
 Arg Gln Ser Asp Arg Gly Thr Arg Leu Ile His Val Ile Asn Ala Leu
 405 410 415
 Pro Pro Thr Cys Pro Pro Gln Lys Lys Lys Lys Lys Lys Lys Gly
 420 425 430

Gly Arg

<210> 139
 <211> 237
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (55)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (237)
 <223> Xaa equals stop translation

<400> 139
 Met Ile Ser Leu Pro Gly Pro Leu Val Thr Asn Leu Leu Arg Phe Leu
 1 5 10 15
 Phe Leu Gly Leu Ser Ala Leu Ala Pro Pro Ser Arg Ala Gln Leu Gln
 20 25 30
 Leu His Leu Pro Ala Asn Arg Leu Gln Ala Val Glu Gly Gly Glu Val
 35 40 45
 Val Leu Pro Ala Trp Tyr Xaa Leu His Gly Glu Val Ser Ser Ser Gln
 50 55 60
 Pro Trp Glu Val Pro Phe Val Met Trp Phe Phe Lys Gln Lys Glu Lys
 65 70 75 80
 Glu Asp Gln Val Leu Ser Tyr Ile Asn Gly Val Thr Thr Ser Lys Pro
 85 90 95

Gly Val Ser Leu Val Tyr Ser Met Pro Ser Arg Asn Leu Ser Leu Arg
 100 105 110
 Leu Glu Gly Leu Gln Glu Lys Asp Ser Gly Pro Tyr Ser Cys Ser Val
 115 120 125
 Asn Val Gln Asp Lys Gln Gly Lys Ser Arg Gly His Ser Ile Lys Thr
 130 135 140
 Leu Glu Leu Asn Val Leu Val Pro Pro Ala Pro Pro Ser Cys Arg Leu
 145 150 155 160
 Gln Gly Val Pro His Val Gly Ala Asn Val Thr Leu Ser Cys Gln Ser
 165 170 175
 Pro Arg Ser Lys Pro Ala Val Gln Tyr Gln Trp Asp Arg Gln Leu Pro
 180 185 190
 Ser Phe Gln Thr Phe Phe Ala Pro Ala Leu Asp Val Ile Arg Gly Ser
 195 200 205
 Leu Ser Leu Thr Asn Leu Ser Ser Ser Met Ala Gly Val Tyr Val Cys
 210 215 220
 Lys Ala His Asn Glu Val Gly Thr Ala Asn Val Met Xaa
 225 230 235

<210> 140
 <211> 100
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (78)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (100)
 <223> Xaa equals stop translation

<400> 140
 Met Thr Trp Gly Thr Trp Leu Val His Thr Phe Leu Cys Ser Val Ala
 1 5 10 15
 Ser Ala Lys Thr Leu Lys Ser Val Arg Lys Tyr Leu Ser Leu Cys Ser
 20 25 30
 Pro Ile Gly Ser Ser Phe Val Val Ser Glu Gly Ser Tyr Leu Asp Ile
 35 40 45
 Ser Asp Trp Leu Asn Pro Ala Lys Leu Ser Leu Tyr Tyr Gln Ile Asn
 50 55 60
 Ala Thr Ser Pro Trp Val Arg Asp Leu Cys Gly Gln Arg Xaa Thr Asp
 65 70 75 80
 Ala Cys Glu Gln Leu Cys Asp Pro Glu Thr Gly Glu Pro Trp Glu Pro
 85 90 95
 Gly Trp Gly Xaa
 100

<210> 141
 <211> 70
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (56)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (70)
 <223> Xaa equals stop translation

<400> 141
 Met Tyr Lys Ala Phe Leu Leu Ala Leu Thr Thr Val Phe Tyr Leu Gly
 1 5 10 15
 Ile Leu Asn Ser His Phe His Gly Cys Val Leu Cys Asn Thr Asn Val
 20 25 30
 Phe Lys Trp Tyr Ser His Pro Val Gly Gln Leu Ser Lys Arg Cys Leu
 35 40 45
 Asp Ala Ser Lys Leu Ala Tyr Xaa Lys Phe Thr Ser Ile Lys Tyr Gln
 50 55 60
 Cys Asn Tyr Ser Thr Xaa
 65 70

<210> 142
 <211> 62
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (62)
 <223> Xaa equals stop translation

<400> 142
 Met His Glu Cys Gln Ser Phe Pro Leu Cys Val His Leu Arg Leu Val
 1 5 10 15
 Leu Leu Leu Ser Phe Lys Thr Gln Val His Glu Phe His Glu Val Phe
 20 25 30
 Pro His Tyr Ser His Phe Asn Phe Pro Ser Leu Asn Asn Tyr Asp Ile
 35 40 45
 Asn Leu Leu Leu Asn His Glu Leu Trp His Thr Thr Pro Xaa
 50 55 60

<210> 143
 <211> 89
 <212> PRT
 <213> Homo sapiens

<220>

<221> SITE
 <222> (73)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (89)
 <223> Xaa equals stop translation

 <400> 143
 Met Asn Leu Val Gly Phe Cys Leu Phe Ile Cys Leu Leu Leu Met Leu
 1 5 10 15

 Leu Leu Leu Leu Leu Phe Ser Lys Phe Ser Ile Val Glu Lys Tyr Ala
 20 25 30

 Ala Pro Glu Glu Met Ile Gly His Ser Pro Ala Trp Cys Trp Thr Leu
 35 40 45

 Ser Ser Leu Ala Gln Pro Ser Pro Asp Leu Ser Val Tyr Leu Thr Leu
 50 55 60

 Val Phe Tyr Ile Leu Gln Arg Gln Xaa Gln Asn Asn Pro Asn Leu Thr
 65 70 75 80

 Gln Ile Pro Gly Ile His Leu Ile Xaa
 85

 <210> 144
 <211> 79
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (40)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (46)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (60)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (79)
 <223> Xaa equals stop translation

 <400> 144
 Met Met Gly Asn Asp Leu Leu His Leu Val Phe Leu Gln Leu Ser Leu
 1 5 10 15

 Gly Val Ala Ser Gly Gly Trp Ile Leu Trp Pro Leu Arg Arg Leu Gly
 20 25 30

 Gly Ala His Thr Ser Lys Asp Xaa Asn Lys Asn Gly His Xaa Val His
 35 40 45

Cys Leu Val Ile Thr Asn Glu Pro Leu Val Ser Xaa Lys Lys Ile Gly
50 55 60

Leu Ser Ser Pro His Thr Cys Pro Ser Thr Leu Gln Gln Phe Xaa
65 70 75

<210> 145
<211> 91
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (91)
<223> Xaa equals stop translation

<400> 145
Met Met Val Trp Asn Leu Phe Pro Cys Phe Pro Pro Leu Leu Leu Leu
1 5 10 15

Gln Phe Ile Asp Cys Gln Gln Ser Ser Glu Ile Glu Gln Gly Phe Thr
20 25 30

Arg Ser Leu Leu Gly His Pro Ile Phe Phe Cys Pro Asp Pro Cys Trp
35 40 45

Gln Ser Cys Met Asn Cys Val Ile Leu Leu Ser Ala Phe Phe Phe Leu
50 55 60

Phe Asp Lys Met Asp Ile Lys Asn Ser Cys Cys Ala Lys Val Ser Ser
65 70 75 80

Leu Leu Gln Glu Glu Asn Gln Phe Phe Phe Xaa
85 90

<210> 146
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

<400> 146
Met Tyr Leu Gly Ser Arg Ile Val Lys Ala Leu Phe Phe Leu Leu Phe
1 5 10 15

Cys Ile Phe His Ile Trp Tyr Asn Glu His Val Leu Arg Thr Val Leu
20 25 30

Asp Leu Arg Lys Tyr Ala Asn Thr Val Gln Ile Val Leu Ala Ser Pro
35 40 45

Met Pro Ser Ser Ser Ile Ala Asn Val Ser Thr Leu Val Trp Cys Val
50 55 60

Cys Cys Asn Gly Xaa
65

<210> 147
 <211> 44
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (44)
 <223> Xaa equals stop translation

<400> 147
 Met Lys Cys Thr Glu Lys Cys Val Val Phe Phe Thr Phe Val Leu
 1 5 10 15
 Tyr Met Tyr Val Tyr Trp Val Leu Trp Ala Val Glu Ala Lys Leu Thr
 20 25 30
 Ser His Val Ala His Glu Met Leu Val Ser Cys Xaa
 35 40

<210> 148
 <211> 85
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (71)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (85)
 <223> Xaa equals stop translation

<400> 148
 Met Gly Cys Ile Pro Leu Ile Lys Ser Ile Ser Asp Trp Arg Val Ile
 1 5 10 15
 Ala Leu Ala Ala Leu Trp Phe Cys Leu Ile Gly Leu Ile Cys Gln Ala
 20 25 30
 Leu Cys Ser Glu Asp Gly His Lys Arg Arg Ile Leu Thr Leu Gly Leu
 35 40 45
 Gly Phe Leu Val Ile Pro Phe Leu Pro Ala Ser Asn Leu Phe Phe Arg
 50 55 60
 Val Gly Phe Val Val Ala Xaa Cys Ser Ser Thr Ser Pro Ala Leu Gly
 65 70 75 80
 Thr Val Cys Cys Xaa
 85

<210> 149
 <211> 64
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (64)

<223> Xaa equals stop translation

<400> 149

```
Met Phe Ile Leu Leu Ile Val Phe Val Phe Ser Lys Ser Lys Gln Val
 1             5             10             15
Leu Ser Ile Cys Leu Lys Ile Phe Lys Val Glu Ile Asn Ser Ile Ser
          20             25             30
Phe Cys Lys Asn Lys Lys Tyr Lys Asp Leu Pro Tyr Ala Phe Ala Ser
          35             40             45
Glu Lys Thr Gly Arg Thr Tyr Ser Asn Val Asn Asn Asp Tyr Leu Xaa
 50             55             60
```

<210> 150

<211> 62

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (62)

<223> Xaa equals stop translation

<400> 150

```
Met Ile Val Tyr Trp Met Ile Trp Ala Leu Arg Ser Pro Leu Thr Thr
 1             5             10             15
Ala Gln Asn Ile His Ser Ser Thr Ala Leu Thr Glu Phe Ala Lys Cys
          20             25             30
Ile Lys Glu Val Thr Trp Arg Val Arg Ser Tyr Glu Thr Ile Cys Arg
          35             40             45
Lys Trp Gly Lys Lys Gly His Met Ala Gln Leu Lys Leu Xaa
 50             55             60
```

<210> 151

<211> 83

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (83)

<223> Xaa equals stop translation

<400> 151

```
Met Arg Phe Phe Leu Glu Cys Val Leu Leu Ile Cys Phe Arg Ala Met
 1             5             10             15
Ser Ala Ile Tyr Thr His Thr Ser Ile Gly Asn Ala Gln Lys Leu Phe
          20             25             30
Thr Asp Gly Ser Ala Phe Arg Arg Val Arg Glu Pro Leu Pro Lys Glu
          35             40             45
Gly Lys Ser Trp Pro Gln Leu Glu Gln Ala Cys Leu Gly Pro Cys Ser
```

50 55 60

Val Phe Gln Leu Gln Thr Ala Cys Ile Ile Pro Ser Cys Tyr Ser Ser
65 70 75 80

Phe Thr Xaa

<210> 152
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 152
Met Cys Cys Ala Ser His Pro Cys Gln Arg Glu Gly Trp Leu Cys Val
1 5 10 15

Ile Phe Thr Val Phe Leu Lys Val Thr Val Cys Val Phe Thr Phe Val
20 25 30

Gln Ile Thr Gly Ser Lys Ala Ala Asn Ser Ala Ile Thr Cys Xaa
35 40 45

<210> 153
<211> 188
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (188)
<223> Xaa equals stop translation

<400> 153
Met Ala Cys Lys Gly Leu Leu Gln Gln Val Gln Gly Pro Arg Leu Pro
1 5 10 15

Trp Thr Arg Leu Leu Leu Leu Leu Val Phe Ala Val Gly Phe Leu
20 25 30

Cys His Asp Leu Pro Val Thr Gln Leu Leu Pro Gly Trp Leu Gly Glu
35 40 45

Thr Leu Pro Leu Trp Gly Ser His Leu Leu Thr Val Val Arg Pro Ser
50 55 60

Leu Gln Leu Ala Trp Ala His Thr Asn Ala Thr Val Ser Phe Leu Ser
65 70 75 80

Ala His Cys Ala Ser His Leu Ala Trp Phe Gly Asp Ser Leu Thr Ser
85 90 95

Leu Ser Gln Arg Leu Gln Ile Gln Leu Pro Asp Ser Val Asn Gln Leu
100 105 110

Leu Arg Tyr Leu Arg Glu Leu Pro Leu Leu Phe His Gln Asn Val Leu
115 120 125

Leu Pro Leu Trp His Leu Leu Leu Glu Ala Leu Ala Trp Ala Gln Glu
 130 135 140

His Cys His Glu Ala Cys Arg Gly Glu Val Thr Trp Asp Cys Met Lys
 145 150 155 160

Thr Gln Leu Ser Glu Ala Val His Trp Thr Trp Leu Cys Tyr Arg Thr
 165 170 175

Leu Gln Trp Leu Ser Trp Thr Gly His Leu Pro Xaa
 180 185

<210> 154
 <211> 114
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (114)
 <223> Xaa equals stop translation

<400> 154
 Met Ile Phe Ser Met Pro Gln Gln Gly Ser Ser Trp Phe Leu Ser Ala
 1 5 10 15

Phe Leu Ser Trp Pro Leu Ala Leu Ala Pro Ala Leu Thr Pro Thr Pro
 20 25 30

Ala Pro Ala Arg Ala Pro Gly Ala Pro Arg Ala Ala Gly Ala Pro Gly
 35 40 45

Arg Val Ala Ala Gly Arg Gly Thr Cys Ala Gly Ala Leu Ala Pro Gly
 50 55 60

Gln Glu Ala Trp Ser Ala Val Trp Glu Pro Gly Leu Phe Ile Trp Val
 65 70 75 80

Glu His Pro Leu Gly Cys Gln Gly His Gly Leu Asp Arg Phe Pro Leu
 85 90 95

Pro Thr Ala Leu Pro Leu Gln Gly Gly His Ala Ala Cys Cys Pro Gln
 100 105 110

Leu Xaa

<210> 155
 <211> 293
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (293)
 <223> Xaa equals stop translation

<400> 155
 Met Gly Ile Gln Thr Ser Pro Val Leu Leu Ala Ser Leu Gly Val Gly
 1 5 10 15

Leu Val Thr Leu Leu Gly Leu Ala Val Gly Ser Tyr Leu Val Arg Arg
 20 25 30
 Ser Arg Arg Pro Gln Val Thr Leu Leu Asp Pro Asn Glu Lys Tyr Leu
 35 40 45
 Leu Arg Leu Leu Asp Lys Thr Thr Val Ser His His Thr Leu Gly Leu
 50 55 60
 Pro Val Gly Lys His Ile Tyr Leu Ser Thr Arg Ile Asp Gly Ser Leu
 65 70 75 80
 Val Ile Arg Pro Tyr Thr Pro Val Thr Ser Asp Glu Asp Gln Gly Tyr
 85 90 95
 Val Asp Leu Val Ile Lys Val Tyr Leu Lys Gly Val His Pro Lys Phe
 100 105 110
 Pro Glu Gly Gly Lys Met Ser Gln Tyr Leu Asp Ser Leu Lys Val Gly
 115 120 125
 Asp Val Val Glu Phe Arg Gly Pro Ser Gly Leu Leu Thr Tyr Thr Gly
 130 135 140
 Lys Gly His Phe Asn Ile Gln Pro Asn Lys Lys Ser Pro Pro Glu Pro
 145 150 155 160
 Arg Val Ala Lys Lys Leu Gly Met Ile Ala Gly Gly Thr Gly Ile Thr
 165 170 175
 Pro Met Leu Gln Leu Ile Arg Ala Ile Leu Lys Val Pro Glu Asp Pro
 180 185 190
 Thr Gln Cys Phe Leu Leu Phe Ala Asn Gln Thr Glu Lys Asp Ile Ile
 195 200 205
 Leu Arg Glu Asp Leu Glu Glu Leu Gln Ala Arg Tyr Pro Asn Arg Phe
 210 215 220
 Lys Leu Trp Phe Thr Leu Asp His Pro Pro Lys Asp Trp Ala Tyr Ser
 225 230 235 240
 Lys Gly Phe Val Thr Ala Asp Met Ile Arg Glu His Leu Pro Ala Pro
 245 250 255
 Gly Asp Asp Val Leu Val Leu Leu Cys Gly Pro Pro Pro Met Val Gln
 260 265 270
 Leu Ala Cys His Pro Asn Leu Asp Lys Leu Gly Tyr Ser Gln Lys Met
 275 280 285
 Arg Phe Thr Tyr Xaa
 290

<210> 156
 <211> 87
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (87)
 <223> Xaa equals stop translation

<400> 156

Met Val Met Val Phe Phe Leu Thr Phe Ser Gly Ser His Gly Cys Val
 1 5 10 15

Pro Thr Ser Gln Pro Trp Lys Asp Ala Glu Asp Gln Val Gly Cys Val
 20 25 30

His Ala Val Ala Trp Val Asn Ser Ala Leu Tyr Thr Val Leu Cys Pro
 35 40 45

Phe Leu Gly Lys Pro Lys Cys Ser Phe Ser Phe Asp Arg Asn Glu Ser
 50 55 60

Glu Asp Leu Asn Lys Gln Glu Val Lys Cys Arg Ala Val Pro Val Ser
 65 70 75 80

Val Ser Ser Ser Met Leu Xaa
 85

<210> 157

<211> 107

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (107)

<223> Xaa equals stop translation

<400> 157

Met Leu Ala Thr Met Val Val Gln Ile Leu Arg Leu Arg Pro His Thr
 1 5 10 15

Gln Lys Trp Ser His Val Leu Thr Leu Leu Gly Leu Ser Leu Val Leu
 20 25 30

Gly Leu Pro Trp Ala Leu Ile Phe Phe Ser Phe Ala Ser Gly Thr Phe
 35 40 45

Gln Leu Val Val Leu Tyr Leu Phe Ser Ile Ile Thr Ser Phe Gln Gly
 50 55 60

Phe Leu Ile Phe Ile Trp Tyr Trp Ser Met Arg Leu Gln Ala Arg Gly
 65 70 75 80

Gly Pro Ser Pro Leu Lys Ser Asn Ser Asp Ser Ala Arg Leu Pro Ile
 85 90 95

Ser Ser Gly Ser Thr Ser Ser Ser Arg Ile Xaa
 100 105

<210> 158

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

100

<400> 158

Met Ala Trp Arg Val Trp Cys Leu Trp Gly Ile Pro Pro Leu Phe Cys
 1 5 10 15

Ser Pro Gly Thr Leu Ser Cys Val Cys Val Ser Phe Leu Ser Pro Gly
 20 25 30

Asn Gly Met Ala Ser Glu His His Pro Arg Ser Ile Phe Pro Leu Gln
 35 40 45

Asn Asp Val Ser Ser His Val Cys Phe Cys Xaa
 50 55

<210> 159

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 159

Met Arg Ser Asp Cys Val Leu Ile Trp Gln Leu Val Gly Val Leu Leu
 1 5 10 15

Ala Ser Gly Leu Ser Gly Asp Arg Ala Pro Leu Ile Val Leu Thr Ala
 20 25 30

Cys Asp Lys Ala Trp Ala Thr Val Xaa
 35 40

<210> 160

<211> 66

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (29)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (35)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (63)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (64)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (66)

<223> Xaa equals stop translation

<400> 160

Met Trp Ala Cys Trp Gly Met Leu Gly Cys Ile Pro Leu Phe Val Pro
 1 5 10 15

Trp Val Pro Val Leu Gly Lys His Phe Ser Gly Cys Xaa Tyr Leu Cys
 20 25 30

Gly Arg Xaa Pro Cys Trp Ile Ala Phe Ile Cys Val Arg Thr Pro Cys
 35 40 45

Gly Pro Thr Thr Ala Pro Thr Ala Thr Leu Lys Trp Ser Pro Xaa Xaa
 50 55 60

Thr Xaa
 65

<210> 161

<211> 47

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals stop translation

<400> 161

Met Arg Tyr Trp Thr Asp Met Arg Arg Asn Tyr Arg Val Thr Tyr Gln
 1 5 10 15

Val Val Leu Leu Phe Leu Cys Phe Ser Leu Leu Thr Glu Cys Lys Thr
 20 25 30

Phe Glu Pro Arg Ser Glu Arg Ser Leu Phe Ser Tyr Pro Leu Xaa
 35 40 45

<210> 162

<211> 141

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (141)

<223> Xaa equals stop translation

<400> 162

Met Phe Ala Gly Leu Phe Phe Leu Phe Phe Val Arg Phe Gly Ile Gly
 1 5 10 15

Arg Gln Leu Leu Ile Lys Phe Pro Trp Phe Phe Ser Phe Gly Tyr Phe
 20 25 30

Ser Lys Gln Gly Pro Thr Gln Lys Gln Ile Asp Ala Ala Ser Phe Thr
 35 40 45

Leu Thr Phe Phe Gly Gln Gly Tyr Ser Gln Gly Thr Gly Thr Asp Lys
 50 55 60

Asn Lys Pro Asn Ile Lys Ile Cys Thr Gln Val Lys Gly Pro Glu Ala
 65 70 75 80

Gly Tyr Val Ala Thr Pro Ile Ala Met Val Gln Ala Ala Met Thr Leu
 85 90 95
 Leu Ser Asp Ala Ser His Leu Pro Lys Ala Gly Gly Val Phe Thr Pro
 100 105 110
 Gly Ala Ala Phe Ser Lys Thr Lys Leu Ile Asp Arg Leu Asn Lys His
 115 120 125
 Gly Ile Glu Phe Ser Val Ile Ser Ser Ser Glu Val Xaa
 130 135 140

<210> 163
 <211> 54
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (54)
 <223> Xaa equals stop translation

<400> 163
 Met Gln Glu Cys Leu Leu His Gly Cys Cys Cys Tyr Leu Leu Arg Leu
 1 5 10 15
 Gly Val Leu Gly Thr Val Gln Cys Ile Ser Thr Trp Leu Ile Leu Thr
 20 25 30
 Ala Asn Glu Gln His Arg Leu Lys Glu Thr Ser Asn Ser Gln Ser Pro
 35 40 45
 Ala Val Ser Arg Ala Xaa
 50

<210> 164
 <211> 168
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (168)
 <223> Xaa equals stop translation

<400> 164
 Met Cys Gly Phe Leu Ser Leu Gln Ile Met Gly Pro Leu Ile Val Leu
 1 5 10 15
 Val Gly Leu Cys Phe Phe Val Val Ala His Val Lys Lys Arg Asn Thr
 20 25 30
 Leu Asn Ala Gly Gln Asp Ala Ser Glu Arg Glu Gly Gln Ile Gln
 35 40 45
 Ile Met Glu Pro Val Gln Val Thr Val Gly Asp Ser Val Ile Ile Phe
 50 55 60
 Pro Pro Pro Pro Pro Pro Tyr Phe Pro Glu Ser Ser Ala Ser Ala Val
 65 70 75 80


```
<210> 165
<211> 63
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> SITE  
<222> (63)  
<223> Xaa equals stop translation
```

```

<400> 165
Met Ser Ile Ser Leu Ser Ser Leu Ile Leu Leu Pro Ile Trp Ile Asn
  1             5             10             15

Met Ala Gln Ile Gln Gln Gly Gly Pro Asp Glu Lys Glu Lys Thr Thr
      20             25             30

Ala Leu Lys Asp Leu Leu Ser Arg Ile Asp Leu Asp Glu Leu Met Lys
      35             40             45

Lys Asp Glu Pro Pro Leu Asp Phe Leu Ile Pro Trp Lys Val Xaa
      50             55             60

```

```
<210> 166
<211> 114
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> SITE  
<222> (114)  
<223> Xaa equals stop translation
```

```

<400> 166
Met His Pro Pro Leu Thr Pro Pro Thr Pro Leu Cys Leu Trp Leu Arg
  1             5             10             15
Leu Leu Lys Ala Gln Ile Leu Ser Tyr Pro Val Pro Arg Phe Glu Thr
      20             25             30
His Ser Leu Ile Ser Arg Cys Ser Gln Val Pro Pro Thr Phe Leu Trp
      35             40             45
Asp Ile Lys Lys Gly Val Arg Gly Gln Arg Glu Pro Ser Gly Pro Leu

```

104

50 55 60
 Leu Pro Tyr Thr Leu His Cys Pro Phe Ser Pro His Gln Asn Ala Gln
 65 70 75 80
 Arg Arg Cys Asp Asp Ala Thr Glu Asp Tyr Ala Thr Trp Ser Asn Arg
 85 90 95
 Ser Gly Gln His Asp Gln Leu Ser Arg Gly Cys Leu Leu Pro Phe Leu
 100 105 110

Leu Xaa

<210> 167
 <211> 62
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (37)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (39)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (62)
 <223> Xaa equals stop translation

<400> 167
 Met Gly Arg Leu Gly Leu Cys Leu Leu Arg Ser Leu Trp Val Pro Gln
 1 5 10 15
 Arg Arg Ala Thr Thr Leu Gly Trp Thr Leu Ala Leu Arg Val Leu Pro
 20 25 30
 Thr Ala Arg Ala Xaa Arg Xaa Leu Pro Val Ala Ala Asp Thr Ala Arg
 35 40 45
 Arg Ala Cys Gly Ala His Thr Arg Ile Arg Val Leu Gly Xaa
 50 55 60

<210> 168
 <211> 42
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (41)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

105

<400> 168

Met Asp Ile Asn Phe Cys Leu Arg Gly Arg His Gly Val Leu Phe Cys
 1 5 10 15

Phe Val Leu Phe Cys Phe Cys His Leu Leu Thr Val Leu Ser Thr His
 20 25 30

Arg Ala Phe Tyr Tyr Leu Ser Ala Xaa Xaa
 35 40

<210> 169

<211> 43

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (43)

<223> Xaa equals stop translation

<400> 169

Met Ile Lys Leu Gln Lys Val Ser Glu Val Ile Lys Val Leu Lys Met
 1 5 10 15

Leu Leu Tyr Pro Leu Val Leu Leu Leu Ser Leu Lys Leu Asp Thr Lys
 20 25 30

Ala Thr Ile Phe Ala Val Leu Glu Asp Val Xaa
 35 40

<210> 170

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 170

Met Tyr Phe Phe Thr Phe Tyr Phe Ser Ile Ser Ser Phe Met Phe Phe
 1 5 10 15

Leu Leu Val Ile Val Lys Ala Thr Asn Gly Pro Arg Tyr Val Val Gly
 20 25 30

Cys Arg Arg Gln Val Ile Leu Tyr Ile Cys Ile Val Pro Asp Asp Xaa
 35 40 45

<210> 171

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 171

Met Ser Gly Phe Lys Glu Phe Asp Phe Val Val Pro Trp Trp Ser Ile
 1 5 10 15
 Ser Phe Leu Leu Ser Phe Leu Leu Leu Leu Ser Phe Trp Ser Leu
 20 25 30
 Trp Val Tyr Thr Phe His Gln Ile Trp Asn Ile Phe Gly Tyr Tyr Phe
 35 40 45
 Ser Lys Xaa
 50

<210> 172

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 172

Met Cys Ile Ser Gly Cys Leu Phe His Cys Ser Ile Cys Leu Phe Phe
 1 5 10 15
 Met Leu Val Pro Tyr Cys Phe Asp Tyr Cys Leu Val Met Tyr Phe Glu
 20 25 30
 Ile Lys Thr Cys Gly Tyr Leu Leu Leu Cys Ser Pro Cys Gln Asp Tyr
 35 40 45
 Ser Arg Ser Phe Val Ala Ser Ser Xaa
 50 55

<210> 173

<211> 170

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (163)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 173

Met Ala Ala Gly Pro Gly Thr His Leu Ser Leu Phe Trp Ala Arg Ile
 1 5 10 15
 Ala Thr Leu Ala Val Trp Ala Ala Ala Ala Leu Val Thr Val Pro Thr
 20 25 30
 Ala Val Phe Gly Val Glu Gly Glu Val Cys Gly Val Arg Leu Cys Leu
 35 40 45
 Leu Arg Phe Pro Ser Arg Tyr Trp Leu Gly Ala Tyr Gln Leu Gln Arg
 50 55 60
 Val Val Leu Ala Phe Met Val Pro Leu Gly Val Ile Thr Thr Ser Tyr

107

65 70 75 80
 Leu Leu Leu Leu Ala Phe Leu Gln Arg Arg Gln Arg Arg Arg Gln Asp
 85 90 95
 Ser Arg Val Val Ala Arg Ser Val Arg Ile Leu Val Ala Ser Phe Phe
 100 105 110
 Leu Cys Trp Phe Pro Asn His Val Val Thr Leu Trp Gly Val Leu Val
 115 120 125
 Lys Phe Asp Leu Val Pro Trp Asn Ser Thr Phe Tyr Thr Ile Gln Thr
 130 135 140
 Tyr Val Phe Pro Val Thr Thr Cys Leu Ala His Ser Asn Ser Cys Leu
 145 150 155 160
 Asn Pro Xaa Ala Tyr Val Leu Ser Arg Ile
 165 170

<210> 174
 <211> 45
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (45)
 <223> Xaa equals stop translation

<400> 174
 Met Phe Arg Ser Ser Ile Ser Leu Met Val Phe Ser Leu Ile Leu Leu
 1 5 10 15
 Leu Thr Thr Glu Arg Arg Ile Leu Ala Cys Pro Pro Ile Ile Leu Asn
 20 25 30
 Ser Ser Ile Phe Leu Ser Asp Leu Ser Val Leu Pro Xaa
 35 40 45

<210> 175
 <211> 47
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (47)
 <223> Xaa equals stop translation

<400> 175
 Met Asn Pro Leu Ser Phe Leu Phe Cys Phe Ile Ile Cys Arg Leu Leu
 1 5 10 15
 Ala Glu Asn Ala Ile Asn Ile Glu Ile Leu Thr Gly Thr Tyr Glu Asn
 20 25 30
 Phe Pro Thr Lys Ala Tyr Tyr Phe Arg Gln Arg Ser Arg Lys Xaa
 35 40 45

<210> 176

108

<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 176
Met Ala Ser Leu Leu Arg Thr Cys Cys Val Pro Tyr Ile Val Leu Ser
1 5 10 15
Ile Tyr Leu Asp Tyr Leu Ile Lys Ser Ser Gln Ser Leu Tyr Leu Thr
20 25 30
Asp Gly Glu Ile Lys Ala His Gly Thr Xaa
35 40

<210> 177
<211> 48
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (48)
<223> Xaa equals stop translation

<400> 177
Met Leu Gln Asp Leu Leu Ser Ala Leu Trp Phe Cys His Pro Cys Cys
1 5 10 15
Leu Cys Cys Gly Leu Cys Trp Leu Gly Val Asp Ala Gly Cys Ser Gln
20 25 30
Gly Gly Ser Gly Cys Pro Gln Gly Lys Ile Ser Asn Asn Gly Ile Xaa
35 40 45

<210> 178
<211> 71
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (71)
<223> Xaa equals stop translation

<400> 178
Met Lys Phe Ala Pro Val Tyr Met Tyr Leu Ser Phe Ile Cys Leu Cys
1 5 10 15
Leu Phe Tyr Cys Asn Ser Ile Asp Thr His His Cys Phe Val Ser Asp
20 25 30
Tyr Leu Ala Phe Glu Ser Ser Met Arg Glu Ala Phe Thr Glu Leu Leu
35 40 45

Ile Leu Ile Lys Gly Glu Ser Asn Val Leu Lys Lys Met Gln Asn His
 50 55 60

His Leu Cys Gln Ser Tyr Xaa
 65 70

<210> 179
 <211> 42
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

<400> 179
 Met Gly Leu Lys Leu Pro Ile Phe Leu Trp Phe Leu Tyr Phe Phe Ile
 1 5 10 15

Pro Leu Ser Ser Cys Tyr Leu Leu Leu Leu Pro His Leu Pro Ser Gly
 20 25 30

Ser Trp Asp Ser Met Leu Ser Phe Pro Xaa
 35 40

<210> 180
 <211> 92
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (18)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 180
 Met Ala Gly Cys Leu Gly Ser Tyr Leu Leu Val Met Ile Leu Ile Leu
 1 5 10 15

Cys Xaa Ala His Phe Phe Ile Cys Gly Asn Glu Asp Asn Arg Val Leu
 20 25 30

Arg Tyr Asn Leu Glu Gln Cys Pro Ser His Ser Lys His Val Ile Asn
 35 40 45

Gly Ser Ser Tyr Cys Tyr Tyr Tyr Tyr Tyr Tyr Tyr Leu Glu Asp Arg
 50 55 60

Gly Ser Val Leu Phe Ile Ile Pro Ser Pro Ala Leu Ser Thr Val Pro
 65 70 75 80

Gly Thr Ile Gln Thr Cys Ile Trp Met Asn Asp Lys
 85 90

<210> 181
 <211> 72
 <212> PRT
 <213> Homo sapiens

<220>

<221> SITE
<222> (72)
<223> Xaa equals stop translation

<400> 181
Met Pro Ala Gly Val Pro Met Ser Thr Tyr Leu Lys Met Phe Ala Ala
1 5 10 15
Ser Leu Leu Ala Met Cys Ala Gly Ala Glu Val Val His Arg Tyr Tyr
20 25 30
Arg Pro Asp Leu Thr Ile Pro Glu Ile Pro Pro Lys Arg Gly Glu Leu
35 40 45
Lys Thr Glu Leu Leu Gly Leu Lys Glu Arg Lys His Lys Pro Gln Val
50 55 60
Ser Gln Gln Glu Glu Leu Lys Xaa
65 70

<210> 182
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (23)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (45)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (67)
<223> Xaa equals stop translation

<400> 182
Met Ala Gly Phe Ala Ser Tyr Pro Trp Ser Asp Phe Pro Trp Cys Trp
1 5 10 15
Val Val Cys Phe Ser Phe Xaa Phe Phe Phe Leu Arg Gln Ser Glu Ser
20 25 30
Leu Ser Gln Lys Lys Arg Gln Val Ala Asp Glu Leu Xaa Phe Gly Gln
35 40 45
Ser Lys Arg Asp Ser Asp Gly Gly Trp Met Leu Arg Ser Ser Ala Gly
50 55 60
Asn Ser Xaa
65

<210> 183
<211> 71
<212> PRT
<213> Homo sapiens

<220>

<221> SITE
 <222> (14)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (21)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (71)
 <223> Xaa equals stop translation

 <400> 183
 Met Gln Pro Ser Tyr Pro Leu Ser Trp Ser Gly Gly Val Xaa Leu Pro
 1 5 10 15
 Cys Leu Ala Ser Xaa Leu Thr Leu Leu Phe Leu Leu Gln Pro Leu Met
 20 25 30
 Leu Pro Leu Gly Gly Ser Gln Thr Gln Leu Gly Asn His Ser Val Val
 35 40 45
 Arg Leu Leu Leu Pro Val Gln Arg Leu Gly Phe Ala Glu Val Pro Pro
 50 55 60
 Leu Glu Val Ala Gln Ser Xaa
 65 70

 <210> 184
 <211> 41
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (41)
 <223> Xaa equals stop translation

 <400> 184
 Met Ile Pro Leu Arg Arg Gly Met Val Gly Gly Leu Leu Leu Leu
 1 5 10 15
 Ala Thr Ala Asn Lys Leu Leu Ala Ala Ser Phe Arg Asp Leu Met Asp
 20 25 30
 Val Leu Thr Cys Pro Arg Pro Arg Xaa
 35 40

 <210> 185
 <211> 67
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (36)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE

<222> (67)

<223> Xaa equals stop translation

<400> 185

Met Gln His Leu Leu His Ser Leu Cys Leu Ser Cys Ser Thr Met
 1 5 10 15

Ala Arg Asn Val Pro Ala Ser Pro Ser Pro Ser Ala Val Ile Val Ser
 20 25 30

Phe Leu Arg Xaa Pro Gln Pro Cys Phe Leu Tyr Ser Leu Gln Asn Cys
 35 40 45

Glu Ser Ile Lys Pro Leu Phe Phe Ile Asn Ser Pro Val Ser Ser Ser
 50 55 60

Ser Leu Xaa
 65

<210> 186

<211> 67

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (67)

<223> Xaa equals stop translation

<400> 186

Met Leu Pro Ser Trp Trp Ala Leu Gly Trp Met Thr Leu Lys Ile Leu
 1 5 10 15

Gln Met Trp Val Gln Ala Cys Thr His Thr Met Glu Tyr Gly His Ser
 20 25 30

Tyr Thr Gly Gly Val Glu Ser Gly Ser Ala Ala Trp His Leu Thr Glu
 35 40 45

Val Gly Pro Lys Arg Thr His Asp Tyr Ala Glu Asn Trp Ile Gly Ser
 50 55 60

Leu Ser Xaa
 65

<210> 187

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 187

Met His Phe Ser Val Ala His Ser Ile Trp Gly Ile Leu Ile Leu Leu
 1 5 10 15

Ser Leu Tyr Glu Gly Val Ile Ser Trp Val Phe Asn Phe Gln Met Phe
 20 25 30

Thr Lys Leu Leu Leu Cys Ala Lys His Tyr Ser His Cys Phe Glu Ser
 35 40 45

Xaa

<210> 188
 <211> 67
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (67)
 <223> Xaa equals stop translation

<400> 188
 Met Ser Leu Ile Leu Leu Gly Ser Pro Ile Ile Pro Leu Trp Ser Tyr
 1 5 10 15

Thr Ser Ala Thr Gln Ala Ala Ala Leu Val Thr Ser His Val Trp Lys
 20 25 30

Pro Ser Leu Glu Ala His Gln Ile Asn Ile Ser Pro Glu Pro Ser Ile
 35 40 45

His Tyr Asp Arg Trp His Thr Gln Ser Asn Cys Ser Leu Ile Asn Ser
 50 55 60

Leu Gln Xaa
 65

<210> 189
 <211> 58
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (58)
 <223> Xaa equals stop translation

<400> 189
 Met Lys Gln Thr Tyr Trp Gln Thr His Ile Leu Leu Val Leu Thr Leu
 1 5 10 15

Tyr Phe Ile Val Leu Ala Tyr Ser Pro Phe Leu Arg Phe Leu Leu Arg
 20 25 30

Asn Ile Gly Thr His Pro Leu Leu Cys Ala Glu Gly Ile Thr Ser Phe
 35 40 45

Phe Leu Ser Tyr Lys Asn Met Leu Tyr Xaa
 50 55

<210> 190
 <211> 53
 <212> PRT
 <213> Homo sapiens

<220>

114

<221> SITE
<222> (53)
<223> Xaa equals stop translation

<400> 190
Met Gly Pro Asn Phe Val Val Leu Cys Leu Asn Leu Leu Gln Asp Thr
1 5 10 15
Leu Ala Tyr Ala Thr Ala Leu Leu Asn Glu Lys Glu Gln Ser Gly Ser
20 25 30
Ser Asn Gly Ser Glu Ser Ser Pro Ala Asn Glu Asn Gly Asp Arg His
35 40 45
Leu Gln Gln Val Xaa
50

<210> 191
<211> 44
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation

<400> 191
Met Ile Val Ile Ala Val Ser Leu Ser Leu Phe Cys Asp Val Val Ser
1 5 10 15
Ser Glu Cys Met Ser Cys Phe Thr Pro Lys Phe Ala Asp Ile Val Ala
20 25 30
Asn Ala Tyr Gln Asn Glu Ser Tyr Ile Phe Ile Xaa
35 40

<210> 192
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (53)
<223> Xaa equals stop translation

<400> 192
Met Leu Leu Pro Val Asn Thr Leu Leu Tyr Ile Leu Leu Thr Pro Leu
1 5 10 15
Cys Phe Phe Tyr Gly Thr Ser Arg Pro Pro Tyr Leu Glu Leu Val Thr
20 25 30
Leu Leu Lys Lys Lys Lys Gln Ser Val Gly Phe Ser Val Cys Ile Leu
35 40 45
Glu Ala Gly Arg Xaa
50

<210> 193

115

<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals stop translation

<400> 193
Met Ile Ile Val Leu Phe Ser Leu Ser Phe Leu Pro Leu Leu Pro Ser
1 5 10 15
Leu Leu Leu Ser Ser Tyr Leu Cys Leu Phe Phe Phe Pro Ser Gln Ser
20 25 30
Pro Ser Ser Phe Phe Phe His Leu Xaa
35 40

<210> 194
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (25)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (72)
<223> Xaa equals stop translation

<400> 194
Met Thr Glu Gly His Val Phe Cys Phe Ala Leu Cys Cys Val Leu Val
1 5 10 15
Phe Leu Ser Met Thr Leu Leu Val Xaa Ser Leu Glu Lys Thr Asn Ala
20 25 30
Gly Gly Val Ile Ala Trp Gly Cys Ile Ser Val Ser Val Gln Thr Gln
35 40 45
Thr Phe Ser Ser Pro Thr Ser Tyr Gln Thr Leu Phe Ile Ala Cys Lys
50 55 60
Leu Trp Asn Pro Arg Lys Leu Xaa
65 70

<210> 195
<211> 60
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (37)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE

<222> (60)

<223> Xaa equals stop translation

<400> 195

Met Ile Gly Leu Thr Ile Ile Ala Cys Phe Ala Val Ile Val Ser Ala
 1 5 10 15

Lys Arg Ala Val Glu Arg His Glu Ser Leu Thr Ser Trp Asn Leu Ala
 20 25 30

Lys Lys Ala Lys Xaa Arg Glu Glu Ala Ala Leu Ala Ala Gln Ala Lys
 35 40 45

Ala Asn Asp Ile Leu Ser Asp Lys Val Phe Thr Xaa
 50 55 60

<210> 196

<211> 81

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (81)

<223> Xaa equals stop translation

<400> 196

Met Leu Thr Gly Ser His Pro Gln Thr His Thr Cys Trp Leu Gly Thr
 1 5 10 15

Arg Leu Trp Val Val Leu Ser Cys Leu Ala Ser Leu Thr Val Ser Asp
 20 25 30

Cys Pro Glu His Gln Val Ser Ser Cys Ile Ser Ser Trp Pro Gly Glu
 35 40 45

His Ser Val Ser Phe Gln Pro Phe Pro Pro Phe Pro His Ser Leu Gly
 50 55 60

Gly Thr Glu Val Gly Val Glu Glu Ser Gln Met Ala Gly Val Gly Ile
 65 70 75 80

Xaa

<210> 197

<211> 79

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (79)

<223> Xaa equals stop translation

<400> 197

Met Leu His Met Phe Leu Leu Leu Leu Tyr Phe Phe Lys Asn Ser Lys
 1 5 10 15

Ser Leu Phe Met Cys His Trp Ile Asn Leu Ser Asp Asn Val Ser His
 20 25 30

Lys Asn Leu Leu Asp Arg Leu Phe Phe Ser Cys Thr Leu Asn Gly Gly
 35 40 45
 Val Glu Val Val Ser Gly Glu Gln Trp Ile Thr Lys Ser Lys Leu Trp Lys
 50 55 60
 Ile Val Lys Arg Met Glu Lys Leu Asn Thr Arg Tyr Gln Lys Xaa
 65 70 75

<210> 198
 <211> 116
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (116)
 <223> Xaa equals stop translation

<400> 198
 Met Cys Met Ser Val Gly Ala His Ile Cys Val Cys Val Cys Met Cys
 1 5 10 15
 Val Leu His Val Cys Gly Glu Val Ser Ser Val Arg Ala Cys Asp Ser
 20 25 30
 Trp Asp Leu His Ser Cys Val Leu Pro Gln Arg Pro Gln Pro Gly Gln
 35 40 45
 Ala Leu Thr Phe Cys Ala Pro Cys Ile Glu Pro Val Cys Cys Gly Cys
 50 55 60
 Leu Trp Pro Pro Met Gly Asn Ser Gly Glu Leu Ala Gly Gly Cys Ala
 65 70 75 80
 Gln Ser Pro Gly Cys Cys Tyr Cys His Ser Ala Gln Leu Gly Gln Ala
 85 90 95
 Val Ala Pro Glu Gly Val Arg Arg Glu Leu Trp Glu His Leu Tyr Ser
 100 105 110
 Val Leu Lys Xaa
 115

<210> 199
 <211> 51
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (51)
 <223> Xaa equals stop translation

<400> 199
 Met Pro Gly Cys Trp Val Leu Glu Leu Val Asp His Trp Leu Ala Ser
 1 5 10 15
 Leu Trp Leu Val Val Ala Val Thr Glu Cys Ala Ala Arg Pro Glu Trp
 20 25 30
 Leu Phe Trp Leu Cys Pro Pro Ser Cys Ser Met Pro Gly Gly Gly Gly

35

40

45

Asp Thr Xaa
50

<210> 200
<211> 58
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (58)
<223> Xaa equals stop translation

<400> 200
Met Lys Phe Tyr Ala Val Leu Leu Ser Ile Cys Leu Leu Leu Ser Cys
1 5 10 15

Trp Cys Ala Cys His Val Arg Asp Cys Asn Leu Ile Cys Leu Phe Ser
20 25 30

Thr Val Lys Ala Ile Thr Arg Glu Leu Leu Gln Leu Pro Ser Tyr Val
35 40 45

Lys Arg Phe Phe Phe Asn Ser Leu Arg Xaa
50 55

<210> 201
<211> 57
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (57)
<223> Xaa equals stop translation

<400> 201
Met Leu Val Ala Pro Phe Asn Leu Leu Phe Glu Met Ala Pro Phe Asn
1 5 10 15

Ile Phe Leu Phe Pro Gln Trp Gly Leu Leu Trp Leu Met Leu Tyr Leu
20 25 30

Leu Tyr Val Phe Gln Ala Ser Leu Arg Thr Pro Glu Leu Thr Trp Glu
35 40 45

Arg Val Arg Ser Gln Val Asp Gln Xaa
50 55

<210> 202
<211> 50
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation

<400> 202

Met Leu Leu Thr Cys Ile Leu Leu His Leu Trp Ile Val Val Asp Ser
 1 5 10 15

Val Ile Tyr Met Lys Pro Thr Ser Arg Asp Gly Cys Leu Leu Ser Ala
 20 25 30

Leu Gln Met Ala Arg Ser Leu Ile Ile Gln Leu Asn His Ser Ser Ser
 35 40 45

Asn Xaa
 50

<210> 203

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (45)

<223> Xaa equals stop translation

<400> 203

Met Pro Leu Cys Gly Leu Tyr Cys Leu Arg Ile Leu Met Phe Pro Leu
 1 5 10 15

Arg Ser Ala Asn Ser Val Pro Leu Gln Cys Leu Pro Pro Ser Ser Leu
 20 25 30

Ala Asn Lys Asp Ser His Phe Arg Ala Pro Arg Lys Xaa
 35 40 45

<210> 204

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (18)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (25)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (45)

<223> Xaa equals stop translation

<400> 204

Met Ser Pro Ser Pro Arg Trp Gly Phe Leu Cys Val Leu Phe Thr Ala
 1 5 10 15

Val Xaa Pro Ala Pro Ser Thr Ala Xaa Val Gln Asp Lys Cys Pro Val
 20 25 30

Asn Thr Trp Glu Ala Met Gln Ala Cys Val His Gly Xaa
 35 40 45

120

<210> 205
 <211> 161
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (136)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (161)
 <223> Xaa equals stop translation

<400> 205
 Met Ala Phe Thr Phe Ala Ala Phe Cys Tyr Met Leu Ser Leu Val Leu
 1 5 10 15
 Cys Ala Ala Leu Ile Phe Phe Ala Ile Trp His Ile Ile Ala Phe Asp
 20 25 30
 Glu Leu Arg Thr Asp Phe Lys Ser Pro Ile Asp Gln Cys Asn Pro Val
 35 40 45
 His Ala Arg Glu Arg Leu Arg Asn Ile Glu Arg Ile Cys Phe Leu Leu
 50 55 60
 Arg Lys Leu Val Leu Pro Glu Tyr Ser Ile His Ser Leu Phe Cys Ile
 65 70 75 80
 Met Phe Leu Cys Ala Gln Glu Trp Leu Thr Leu Gly Leu Asn Val Pro
 85 90 95
 Leu Leu Phe Tyr His Phe Trp Arg Tyr Phe His Cys Pro Ala Asp Ser
 100 105 110
 Ser Glu Leu Ala Tyr Asp Pro Pro Val Val Met Asn Ala Asp Thr Leu
 115 120 125
 Ser Tyr Cys Gln Lys Glu Ala Xaa Cys Lys Leu Ala Phe Tyr Leu Leu
 130 135 140
 Ser Phe Phe Tyr Tyr Leu Tyr Cys Met Ile Tyr Thr Leu Val Ser Ser
 145 150 155 160

Xaa

<210> 206
 <211> 198
 <212> PRT
 <213> Homo sapiens

<400> 206
 Met Tyr Arg Glu Arg Leu Arg Thr Leu Leu Val Ile Ala Val Val Met
 1 5 10 15
 Ser Leu Leu Asn Ala Leu Ser Thr Ser Gly Gly Ser Ile Ser Trp Asn
 20 25 30

121

Asp Phe Val His Glu Met Leu Ala Lys Gly Glu Val Gln Arg Val Gln
 35 40 45
 Val Val Pro Glu Ser Asp Val Val Glu Val Tyr Leu His Pro Gly Ala
 50 55 60
 Val Val Phe Gly Arg Pro Arg Leu Ala Leu Met Tyr Arg Met Gln Val
 65 70 75 80
 Ala Asn Ile Asp Lys Phe Glu Glu Lys Leu Arg Ala Ala Glu Asp Glu
 85 90 95
 Leu Asn Ile Glu Ala Lys Asp Arg Ile Pro Val Ser Tyr Lys Arg Thr
 100 105 110
 Gly Phe Phe Gly Lys Cys Pro Val Leu Cys Gly Asp Asp Gly Ser Gly
 115 120 125
 Pro Gly His Pro Val Val Cys Phe Pro Ser Gly Arg Asp Asp Trp Arg
 130 135 140
 His Arg Arg Arg Trp Thr Ser Arg Ser Arg Leu Leu Cys Trp Lys Ala
 145 150 155 160
 Leu Met Gly Ser Val Gly Ala Asp His Thr Arg Glu Leu Arg Lys Pro
 165 170 175
 Ser Gly Ser His Arg Pro Pro Phe Asn Val Val Ile Pro Trp Trp Trp
 180 185 190
 Lys Gln Asp Asp Gly Pro
 195

<210> 207
 <211> 60
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (60)
 <223> Xaa equals stop translation

<400> 207
 Met Asn Ser Thr Leu Cys Val Val Leu Ser Leu Met Cys Met Asn Ser
 1 5 10 15
 Thr Leu Cys Val Val Leu Ser Leu Thr His Ser Cys Pro Ser Pro Gln
 20 25 30
 Val Pro Lys Val His Tyr Met Ile Phe Met Pro Leu His Leu His Ser
 35 40 45
 Leu Ala Leu Thr Gln Leu Ile Ile Ile Tyr Lys Xaa
 50 55 60

<210> 208
 <211> 82
 <212> PRT
 <213> Homo sapiens
 <220>

<221> SITE

<222> (82)

<223> Xaa equals stop translation

<400> 208

Met Val Val Ala Gly Val Val Val Leu Ile Leu Ala Leu Val Leu Ala
 1 5 10 15

Trp Leu Ser Thr Tyr Val Ala Asp Ser Gly Ser Asn Gln Leu Leu Gly
 20 25 30

Ala Ile Val Ser Ala Gly Asp Thr Ser Val Leu His Leu Gly His Val
 35 40 45

Asp His Leu Val Ala Gly Gln Gly Asn Pro Glu Pro Thr Glu Leu Pro
 50 55 60

His Pro Ser Glu Asp Lys Gln Val Gln Ala Ala Val Gln Arg Pro
 65 70 75 80

Pro Xaa

<210> 209

<211> 43

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (18)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (37)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (43)

<223> Xaa equals stop translation

<400> 209

Met Ala Gly Cys Leu Gly Ser Tyr Leu Leu Val Met Ile Leu Ile Leu
 1 5 10 15

Cys Xaa Ala His Phe Phe Ile Cys Gly Asn Glu Asp Asn Arg Val Leu
 20 25 30

Arg Tyr Asn Leu Xaa Thr Met Ser Val Thr Xaa
 35 40

<210> 210

<211> 97

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (97)

<223> Xaa equals stop translation

123

<400> 210

Met Tyr Arg Glu Arg Leu Arg Thr Leu Leu Val Ile Ala Val Val Met
 1 5 10 15
 Ser Leu Leu Asn Ala Leu Ser Thr Ser Gly Gly Ser Ile Ser Trp Asn
 20 25 30
 Asp Phe Val His Glu Met Leu Ala Lys Gly Glu Val Gln Arg Val Gln
 35 40 45
 Val Val Pro Glu Ser Asp Val Val Glu Val Tyr Leu His Pro Gly Ala
 50 55 60
 Val Val Phe Gly Arg Pro Arg Leu Ala Leu Met Tyr Arg Met Gln Leu
 65 70 75 80
 Gln Ile Leu Thr Ser Leu Lys Arg Ser Phe Glu Gln Leu Lys Met Ser
 85 90 95

Xaa

<210> 211

<211> 22

<212> PRT

<213> Homo sapiens

<400> 211

Trp Ala Gly Thr Gln Glu Pro Thr Gly Leu Pro Ser Thr Leu Ser Arg
 1 5 10 15

Ser Glu Ser Trp Asp His
 20

<210> 212

<211> 171

<212> PRT

<213> Homo sapiens

<400> 212

Glu Ile Ile His Asn Leu Pro Thr Ser Arg Met Ala Ala Arg Thr Lys
 1 5 10 15

Lys Lys Asn Asp Ile Ile Asn Ile Lys Val Pro Ala Asp Cys Asn Thr
 20 25 30

Arg Met Ser Tyr Tyr Tyr Lys Gly Ser Gly Lys Arg Gly Glu Met Glu
 35 40 45

Ser Trp Leu Val Met Ser Ser Trp Ser Ile Leu Asp Phe Glu Phe Leu
 50 55 60

Glu Ala Arg Pro Gln Leu Phe Asn Leu Val Tyr Thr Glu His Ser Thr
 65 70 75 80

Tyr Ser Gly Arg His Tyr Thr Arg Glu Arg Gly Gly Phe Met Val Phe
 85 90 95

Lys Asn Ser Tyr Ser Gln Leu Leu Leu Lys Arg Lys Asp Ser Leu Cys
 100 105 110

124

Ala Phe Ile Gln Pro Met Ala Leu Asn Ile Ile His Val Pro Met Ser
115 120 125
Ser Lys Cys Ile Phe Pro Ala Gln Ser Gly Pro Ser Thr Phe Arg Ser
130 135 140
Leu Trp Trp Cys Pro His Pro Ile Ser Lys Cys Gln Leu Gly Leu Tyr
145 150 155 160
Ser Ser Gln Ile Arg Asp Ile Pro Tyr Leu Ala
165 170

<210> 213
<211> 35
<212> PRT
<213> Homo sapiens

<400> 213
Glu Ile Ile His Asn Leu Pro Thr Ser Arg Met Ala Ala Arg Thr Lys
1 5 10 15
Lys Lys Asn Asp Ile Ile Asn Ile Lys Val Pro Ala Asp Cys Asn Thr
20 25 30
Arg Met Ser
35

<210> 214
<211> 36
<212> PRT
<213> Homo sapiens

<400> 214
Tyr Tyr Tyr Lys Gly Ser Gly Lys Arg Gly Glu Met Glu Ser Trp Leu
1 5 10 15
Val Met Ser Ser Trp Ser Ile Leu Asp Phe Glu Phe Leu Glu Ala Arg
20 25 30
Pro Gln Leu Phe
35

<210> 215
<211> 36
<212> PRT
<213> Homo sapiens

<400> 215
Asn Leu Val Tyr Thr Glu His Ser Thr Tyr Ser Gly Arg His Tyr Thr
1 5 10 15
Arg Glu Arg Gly Gly Phe Met Val Phe Lys Asn Ser Tyr Ser Gln Leu
20 25 30
Leu Leu Lys Arg
35

<210> 216
<211> 35
<212> PRT

<213> Homo sapiens

<400> 216

Lys Asp Ser Leu Cys Ala Phe Ile Gln Pro Met Ala Leu Asn Ile Ile
 1 5 10 15

His Val Pro Met Ser Ser Lys Cys Ile Phe Pro Ala Gln Ser Gly Pro
 20 25 30

Ser Thr Phe
 35

<210> 217

<211> 29

<212> PRT

<213> Homo sapiens

<400> 217

Arg Ser Leu Trp Trp Cys Pro His Pro Ile Ser Lys Cys Gln Leu Gly
 1 5 10 15

Leu Tyr Ser Ser Gln Ile Arg Asp Ile Pro Tyr Leu Ala
 20 25

<210> 218

<211> 460

<212> PRT

<213> Homo sapiens

<400> 218

Met Phe Thr Ile Lys Leu Leu Leu Phe Ile Val Pro Leu Val Ile Ser
 1 5 10 15

Ser Arg Ile Asp Gln Asp Asn Ser Ser Phe Asp Ser Leu Ser Pro Glu
 20 25 30

Pro Lys Ser Arg Phe Ala Met Leu Asp Asp Val Lys Ile Leu Ala Asn
 35 40 45

Gly Leu Leu Gln Leu Gly His Gly Leu Lys Asp Phe Val His Lys Thr
 50 55 60

Lys Gly Gln Ile Asn Asp Ile Phe Gln Lys Leu Asn Ile Phe Asp Gln
 65 70 75 80

Ser Phe Tyr Asp Leu Ser Leu Gln Thr Ser Glu Ile Lys Glu Glu Glu
 85 90 95

Lys Glu Leu Arg Arg Thr Thr Tyr Lys Leu Gln Val Lys Asn Glu Glu
 100 105 110

Val Lys Asn Met Ser Leu Glu Leu Asn Ser Lys Leu Glu Ser Leu Leu
 115 120 125

Glu Glu Lys Ile Leu Leu Gln Gln Lys Val Lys Tyr Leu Glu Glu Gln
 130 135 140

Leu Thr Asn Leu Ile Gln Asn Gln Pro Glu Thr Pro Glu His Pro Glu
 145 150 155 160

Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser Ile Lys
 165 170 175

126

Asp Leu Leu Gln Thr Val Glu Asp Gln Tyr Lys Gln Leu Asn Gln Gln
 180 185 190
 His Ser Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser Ile
 195 200 205
 Gln Glu Pro Thr Glu Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro Arg
 210 215 220
 Thr Thr Pro Phe Leu Gln Leu Asn Glu Ile Arg Asn Val Lys His Asp
 225 230 235 240
 Gly Ile Pro Ala Glu Cys Thr Thr Ile Tyr Asn Arg Gly Glu His Thr
 245 250 255
 Ser Gly Met Tyr Ala Ile Arg Pro Ser Asn Ser Gln Val Phe His Val
 260 265 270
 Tyr Cys Asp Val Ile Ser Gly Ser Pro Trp Thr Leu Ile Gln His Arg
 275 280 285
 Ile Asp Gly Ser Gln Asn Phe Asn Glu Thr Trp Glu Asn Tyr Lys Tyr
 290 295 300
 Gly Phe Gly Arg Leu Asp Gly Glu Phe Trp Leu Gly Leu Glu Lys Ile
 305 310 315 320
 Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu Glu
 325 330 335
 Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu Gly
 340 345 350
 Asn His Glu Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly Asn
 355 360 365
 Val Pro Asn Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr Trp
 370 375 380
 Asp His Lys Ala Lys Gly His Phe Asn Cys Pro Glu Gly Tyr Ser Gly
 385 390 395 400
 Gly Trp Trp Trp His Asp Glu Cys Gly Glu Asn Asn Leu Asn Gly Lys
 405 410 415
 Tyr Asn Lys Pro Arg Ala Lys Ser Lys Pro Glu Arg Arg Arg Gly Leu
 420 425 430
 Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr Ser Ile Lys Ser Thr Lys
 435 440 445
 Met Leu Ile His Pro Thr Asp Ser Ser Phe Glu
 450 455 460

<210> 219

<211> 37

<212> PRT

<213> Homo sapiens

<400> 219

Met Phe Thr Ile Lys Leu Leu Leu Phe Ile Val Pro Leu Val Ile Ser
 1 5 10 15

Ser Arg Ile Asp Gln Asp Asn Ser Ser Phe Asp Ser Leu Ser Pro Glu
 20 25 30

Pro Lys Ser Arg Phe
 35

<210> 220
 <211> 34
 <212> PRT
 <213> Homo sapiens

<400> 220
 Ala Met Leu Asp Asp Val Lys Ile Leu Ala Asn Gly Leu Leu Gln Leu
 1 5 10 15

Gly His Gly Leu Lys Asp Phe Val His Lys Thr Lys Gly Gln Ile Asn
 20 25 30

Asp Ile

<210> 221
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 221
 Phe Gln Lys Leu Asn Ile Phe Asp Gln Ser Phe Tyr Asp Leu Ser Leu
 1 5 10 15

Gln Thr Ser Glu Ile Lys Glu Glu Glu Lys Glu Leu Arg Arg Thr Thr
 20 25 30

Tyr Lys Leu
 35

<210> 222
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 222
 Gln Val Lys Asn Glu Glu Val Lys Asn Met Ser Leu Glu Leu Asn Ser
 1 5 10 15

Lys Leu Glu Ser Leu Leu Glu Glu Lys Ile Leu Leu Gln Gln Lys Val
 20 25 30

Lys Tyr Leu Glu
 35

<210> 223
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 223
 Glu Gln Leu Thr Asn Leu Ile Gln Asn Gln Pro Glu Thr Pro Glu His
 1 5 10 15

Pro Glu Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser
 20 25 30

Ile Lys Asp Leu
 35

<210> 224
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 224
 Leu Gln Thr Val Glu Asp Gln Tyr Lys Gln Leu Asn Gln Gln His Ser
 1 5 10 15

Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser Ile Gln Glu
 20 25 30

Pro Thr Glu
 35

<210> 225
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 225
 Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro Arg Thr Thr Pro Phe Leu
 1 5 10 15

Gln Leu Asn Glu Ile Arg Asn Val Lys His Asp Gly Ile Pro Ala Glu
 20 25 30

Cys Thr Thr
 35

<210> 226
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 226
 Ile Tyr Asn Arg Gly Glu His Thr Ser Gly Met Tyr Ala Ile Arg Pro
 1 5 10 15

Ser Asn Ser Gln Val Phe His Val Tyr Cys Asp Val Ile Ser Gly Ser
 20 25 30

Pro Trp Thr Leu
 35

<210> 227
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 227
 Ile Gln His Arg Ile Asp Gly Ser Gln Asn Phe Asn Glu Thr Trp Glu
 1 5 10 15

129

Asn Tyr Lys Tyr Gly Phe Gly Arg Leu Asp Gly Glu Phe Trp Leu Gly
 20 25 30

Leu Glu Lys Ile
 35

<210> 228
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 228
 Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu Glu
 1 5 10 15

Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu Gly
 20 25 30

Asn His Glu
 35

<210> 229
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 229
 Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly Asn Val Pro Asn
 1 5 10 15

Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr Trp Asp His Lys
 20 25 30

Ala Lys Gly
 35

<210> 230
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 230
 His Phe Asn Cys Pro Glu Gly Tyr Ser Gly Gly Trp Trp Trp His Asp
 1 5 10 15

Glu Cys Gly Glu Asn Asn Leu Asn Gly Lys Tyr Asn Lys Pro Arg Ala
 20 25 30

Lys Ser Lys Pro
 35

<210> 231
 <211> 34
 <212> PRT
 <213> Homo sapiens

<400> 231
 Glu Arg Arg Arg Gly Leu Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr
 1 5 10 15

Ser Ile Lys Ser Thr Lys Met Leu Ile His Pro Thr Asp Ser Glu Ser
 20 25 30

Phe Glu

<210> 232
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 232
 Leu Pro Pro Arg Gly Pro Ala Thr Phe Gly Ser Pro Gly Cys Pro Pro
 1 5 10 15

Ala Asn Ser Pro Pro Ser Ala Pro Ala Thr Pro Glu Pro Ala Arg Ala
 20 25 30

Pro Glu Arg Val
 35

<210> 233
 <211> 44
 <212> PRT
 <213> Homo sapiens

<400> 233
 Gly Thr Arg Ala Gly Val Ser Lys Tyr Thr Gly Gly Arg Gly Val Thr
 1 5 10 15

Trp Ala Pro Ser Ser Ala Ala Val Pro Arg Ile Ser Ser Ala Thr Met
 20 25 30

Arg Met Gly Leu Thr Ser Phe Ser Thr Thr Gly Ala
 35 40

<210> 234
 <211> 306
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (293)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 234
 Trp Gln Ser Gly His Arg Leu Trp Gln Leu Glu Trp Pro Pro Pro Pro
 1 5 10 15

Leu Ser Ala Asp Glu His Pro Trp Glu Gly Pro Leu Pro Gly Thr Ser
 20 25 30

Pro Ser Pro Lys Phe Ser Met Pro Ser Pro Val Pro His Gly His His
 35 40 45

Arg Pro Thr Leu Thr Met Thr Arg Ser Trp Arg Ile Phe Phe Asn Asn
 50 55 60

Ile Ala Tyr Arg Ser Ser Ser Ala Asn Arg Leu Phe Arg Val Ile Arg

131

65 70 75 80
 Arg Glu His Gly Asp Pro Leu Ile Glu Glu Leu Asn Pro Gly Asp Ala
 85 90 95
 Leu Glu Pro Glu Gly Arg Gly Thr Gly Gly Val Val Thr Asp Phe Asp
 100 105 110
 Gly Asp Gly Met Leu Asp Leu Ile Leu Ser His Gly Glu Ser Met Ala
 115 120 125
 Gln Pro Leu Ser Val Phe Arg Gly Asn Gln Gly Phe Asn Asn Asn Trp
 130 135 140
 Leu Arg Val Val Pro Arg Thr Arg Phe Gly Ala Phe Ala Arg Gly Ala
 145 150 155 160
 Lys Val Val Leu Tyr Thr Lys Lys Ser Gly Ala His Leu Arg Ile Ile
 165 170 175
 Asp Gly Gly Ser Gly Tyr Leu Cys Glu Met Glu Pro Val Ala His Phe
 180 185 190
 Gly Leu Gly Lys Asp Glu Ala Ser Ser Val Glu Val Thr Trp Pro Asp
 195 200 205
 Gly Lys Met Val Ser Arg Asn Val Ala Ser Gly Glu Met Asn Ser Val
 210 215 220
 Leu Glu Ile Leu Tyr Pro Arg Asp Glu Asp Thr Leu Gln Asp Pro Ala
 225 230 235 240
 Pro Leu Glu Cys Gly Gln Gly Phe Ser Gln Gln Glu Asn Gly His Cys
 245 250 255
 Met Asp Thr Asn Glu Cys Ile Gln Phe Pro Phe Val Cys Pro Arg Asp
 260 265 270
 Lys Pro Val Cys Val Asn Thr Tyr Gly Ser Tyr Arg Cys Arg Thr Asn
 275 280 285
 Lys Lys Cys Ser Xaa Gly Leu Arg Val Pro Thr Arg Met Ala His Thr
 290 295 300
 Gly Leu
 305

 <210> 235
 <211> 36
 <212> PRT
 <213> Homo sapiens

 <400> 235
 Trp Gln Ser Gly His Arg Leu Trp Gln Leu Glu Trp Pro Pro Pro Pro
 1 5 10 15
 Leu Ser Ala Asp Glu His Pro Trp Glu Gly Pro Leu Pro Gly Thr Ser
 20 25 30
 Pro Ser Pro Lys
 35

132

<210> 236
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 236
 Phe Ser Met Pro Ser Pro Val Pro His Gly His His Arg Pro Thr Leu
 1 5 10 15
 Thr Met Thr Arg Ser Trp Arg Ile Phe Phe Asn Asn Ile Ala Tyr Arg
 20 25 30
 Ser Ser Ser
 35

<210> 237
 <211> 37
 <212> PRT
 <213> Homo sapiens

<400> 237
 Ala Asn Arg Leu Phe Arg Val Ile Arg Arg Glu His Gly Asp Pro Leu
 1 5 10 15
 Ile Glu Glu Leu Asn Pro Gly Asp Ala Leu Glu Pro Glu Gly Arg Gly
 20 25 30
 Thr Gly Gly Val Val
 35

<210> 238
 <211> 34
 <212> PRT
 <213> Homo sapiens

<400> 238
 Thr Asp Phe Asp Gly Asp Gly Met Leu Asp Leu Ile Leu Ser His Gly
 1 5 10 15
 Glu Ser Met Ala Gln Pro Leu Ser Val Phe Arg Gly Asn Gln Gly Phe
 20 25 30
 Asn Asn

<210> 239
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 239
 Asn Trp Leu Arg Val Val Pro Arg Thr Arg Phe Gly Ala Phe Ala Arg
 1 5 10 15
 Gly Ala Lys Val Val Leu Tyr Thr Lys Lys Ser Gly Ala His Leu Arg
 20 25 30
 Ile Ile Asp
 35

133

<210> 240
<211> 36
<212> PRT
<213> Homo sapiens

<400> 240
Gly Gly Ser Gly Tyr Leu Cys Glu Met Glu Pro Val Ala His Phe Gly
1 5 10 15
Leu Gly Lys Asp Glu Ala Ser Ser Val Glu Val Thr Trp Pro Asp Gly
20 25 30
Lys Met Val Ser
35

<210> 241
<211> 35
<212> PRT
<213> Homo sapiens

<400> 241
Arg Asn Val Ala Ser Gly Glu Met Asn Ser Val Leu Glu Ile Leu Tyr
1 5 10 15
Pro Arg Asp Glu Asp Thr Leu Gln Asp Pro Ala Pro Leu Glu Cys Gly
20 25 30
Gln Gly Phe
35

<210> 242
<211> 36
<212> PRT
<213> Homo sapiens

<400> 242
Ser Gln Gln Glu Asn Gly His Cys Met Asp Thr Asn Glu Cys Ile Gln
1 5 10 15
Phe Pro Phe Val Cys Pro Arg Asp Lys Pro Val Cys Val Asn Thr Tyr
20 25 30
Gly Ser Tyr Arg
35

<210> 243
<211> 22
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (9)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 243
Cys Arg Thr Asn Lys Lys Cys Ser Xaa Gly Leu Arg Val Pro Thr Arg
1 5 10 15
Met Ala His Thr Gly Leu
20

<210> 244
<211> 9
<212> PRT
<213> Homo sapiens

<400> 244
Gln Ser Pro Ile Asp Ile Gln Thr Asp
1 5

<210> 245
<211> 18
<212> PRT
<213> Homo sapiens

<400> 245
Leu His Asn Asn Gly His Thr Val Gln Leu Ser Leu Pro Ser Thr Leu
1 5 10 15

Tyr Leu

<210> 246
<211> 11
<212> PRT
<213> Homo sapiens

<400> 246
Tyr Val Ala Ala Gln Leu His Leu His Trp Gly
1 5 10

<210> 247
<211> 11
<212> PRT
<213> Homo sapiens

<400> 247
Ala Glu Leu His Ile Val His Tyr Asp Ser Asp
1 5 10

<210> 248
<211> 16
<212> PRT
<213> Homo sapiens

<400> 248
Gly Gln His Trp Thr Tyr Glu Gly Pro His Gly Gln Asp His Trp Pro
1 5 10 15

<210> 249
<211> 14
<212> PRT
<213> Homo sapiens

<400> 249

135

Gln Ser Pro Ile Asp Ile Gln Thr Asp Ser Val Thr Phe Asp
 1 5 10

<210> 250
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 250
 Leu His Asn Asn Gly His Thr Val Gln Leu Ser Leu Pro Ser Thr
 1 5 10 15

<210> 251
 <211> 12
 <212> PRT
 <213> Homo sapiens

<400> 251
 Lys Tyr Val Ala Ala Gln Leu His Leu His Trp Gly
 1 5 10

<210> 252
 <211> 13
 <212> PRT
 <213> Homo sapiens

<400> 252
 Ala Glu Leu His Ile Val His Tyr Asp Ser Asp Ser Tyr
 1 5 10

<210> 253
 <211> 1667
 <212> DNA
 <213> Homo sapiens

<400> 253
 GGCCGCGCCG CCGCTGCCGC CGCCGCGCGC GATTCTGCTT CTCAGAAGAT GCACTATTAT 60
 AGATACTCTA ACGCCAAGGT CAGCTGCTGG TACAAGTACC TCCTTTTCAG CTACAACATC 120
 ATCTTCTGAT TGGCTGGAGT TGTCTTCCTT GGAGTCGGGC TGTGGGCATG GAGCGAAAAG 180
 GGTGTGCTGT CCGACCTCAC CAAAGTGACC CGGATGCATG GAATCGACCC TGTGGTGCTG 240
 GTCCTGATGG TGGGCGTGGT GATGTTACCC CTGGGGTTCG CCGGCTGCGT GGGGGCTCTG 300
 CGGGAGAATA TCTGCTTGCT CAACTTTTTC TGTGGCACCA TCGTGCTCAT CTTCTTCCTG 360
 GAGCTGGCTG TGGCCGTGCT GGCCTTCCTG TTCCAGGACT GGGTGAGGGA CCGGTTCCGG 420
 GAGTTCTTCG AGAGCAACAT CAAGTCCTAC CGGGACGATA TCGATCTGCA AAACCTCATC 480
 GACTCCCTTC AGAAAGCTAA CCAGTGCTGT GGCGCATATG GCCCTGAAAG ACTGGGACCT 540
 CAGACGTCTA CTCAATTGC AGCGGTGCCA GCTACAGCCG AGAGAATGCG GGGTCCCCTT 600
 CTCCTGCTGC GTGCCAGATC CTGCGCAAAA AGTTGTGAAC ACACAGTGTG GATATGATGT 660
 CAGGATTCAG CTGAAGAGCA AGTGGGATGA GTCCATCTTC ACGAAAGGCT GCATCCAGGC 720

GCTGGAAAGC TGGCTCCCGC GGAACATTTA CATTGTGGCT GCGCTCTTCA TCGCCATCTC 780
 GCTGTTGCAG ATATTGGCA TCTTCCTGGC AAGGACGCTG ATCTCAGACA TCGAGGCAGT 840
 GAAGGCCGGC CATCACTTCT GAGGAGCAGA GTTGAGGGAG CCGAGCTGAG CCACGCTGGG 900
 AGGCCAGAGC CTTTCTCTGC CATCAGCCCT ACGTCCAGAG GGAGAGGAGC CGACACCCCC 960
 AGAGCCAGTG CCCCATCTTA AGCATCAGCG TGACGTGACC TCTCTGTTTC TGCTTGCTGG 1020
 TGCTGAAGAC CAAGGGTCCC CCTTGTTACC TGCCCAAAC TGTGACTGCA TCCCTCTGGA 1080
 GTCTACCCAG AGACAGAGAA TGTGTCTTTA TGTGGGAGTG GTGACTCTGA AAGACAGAGA 1140
 GGGCTCCTGT GGCTGCCAGG AGGGCTTGAC TCAGACCCCC TGCAGCTCAA GCATGTCTGC 1200
 AGGACACCTG GTCCCCCTCT CCCAGTGCCA TCCCAAACAT CTGCTTTGGG TCCATCCCAC 1260
 ATCTGTGGGT GGGCCCGTGG GTAAGAAGG AACCACACAG GCGTGAACA GGGCATCCTC 1320
 TCTCCCATCC AAGCAAAGCC AGCATGGGG CCTGCCCGTA ACGGGAGGCG GACGTGGCCC 1380
 CGCTGGGCCT CTGAGTGCCA GCGCAGTCTG CTGGGACATG CACATATCAG GGGTTGTTTG 1440
 CAGGATCCTC AGCCATGTTT AAGTGAAGTA AGCCTGAGCC AGTGCGTGGA CTGGTGCCAC 1500
 GGGAGTGCCT TGTCCACTGT CCCCTGTGT CCACCAGCTA TTCTCCTGGC GCCGGAAGTG 1560
 CCTCTGGTCT TGATAGCATT AAGCCCTGAT TGCCCGTGG CGCGGTGGGC ATGGTTCTTC 1620
 ACTGAGAGCC GGCTCTCCTT TTCTTAAAGT GTGTAAATAG TTTATTT 1667

<210> 254

<211> 270

<212> PRT

<213> Homo sapiens

<400> 254

Met His Tyr Tyr Arg Tyr Ser Asn Ala Lys Val Ser Cys Trp Tyr Lys
 1 5 10 15
 Tyr Leu Leu Phe Ser Tyr Asn Ile Ile Phe Trp Leu Ala Gly Val Val
 20 25 30
 Phe Leu Gly Val Gly Leu Trp Ala Trp Ser Glu Lys Gly Val Leu Ser
 35 40 45
 Asp Leu Thr Lys Val Thr Arg Met His Gly Ile Asp Pro Val Val Leu
 50 55 60
 Val Leu Met Val Gly Val Val Met Phe Thr Leu Gly Phe Ala Gly Cys
 65 70 75 80
 Val Gly Ala Leu Arg Glu Asn Ile Cys Leu Leu Asn Phe Phe Cys Gly
 85 90 95
 Thr Ile Val Leu Ile Phe Phe Leu Glu Leu Ala Val Ala Val Leu Ala
 100 105 110
 Phe Leu Phe Gln Asp Trp Val Arg Asp Arg Phe Arg Glu Phe Phe Glu
 115 120 125

137

Ser Asn Ile Lys Ser Tyr Arg Asp Asp Ile Asp Leu Gln Asn Leu Ile
130 135 140

Asp Ser Leu Gln Lys Ala Asn Gln Cys Cys Gly Ala Tyr Gly Pro Glu
145 150 155 160

Asp Trp Asp Leu Asn Val Tyr Phe Asn Cys Ser Gly Ala Ser Tyr Ser
165 170 175

Arg Glu Lys Cys Gly Val Pro Phe Ser Cys Cys Val Pro Asp Pro Ala
180 185 190

Gln Lys Val Val Asn Thr Gln Cys Gly Tyr Asp Val Arg Ile Gln Leu
195 200 205

Lys Ser Lys Trp Asp Glu Ser Ile Phe Thr Lys Gly Cys Ile Gln Ala
210 215 220

Leu Glu Ser Trp Leu Pro Arg Asn Ile Tyr Ile Val Ala Gly Val Phe
225 230 235 240

Ile Ala Ile Ser Leu Leu Gln Ile Phe Gly Ile Phe Leu Ala Arg Thr
245 250 255

Leu Ile Ser Asp Ile Glu Ala Val Lys Ala Gly His His Phe
260 265 270

<210> 255

<211> 277

<212> PRT

<213> Homo sapiens

<400> 255

Ser Gly Asn Leu Gly Ser Ala Asp Gly Trp Ala Tyr Ile Asp Val Glu
1 5 10 15

Val Arg Arg Pro Trp Ala Phe Val Gly Pro Gly Cys Ser Arg Ser Ser
20 25 30

Gly Asn Gly Ser Thr Ala Tyr Gly Leu Val Gly Ser Pro Arg Trp Leu
35 40 45

Ser Pro Phe His Thr Gly Gly Ala Val Ser Leu Pro Arg Arg Pro Arg
50 55 60

Gly Pro Gly Pro Val Leu Gly Val Ala Arg Pro Cys Leu Arg Cys Val
65 70 75 80

Leu Arg Pro Glu His Tyr Glu Pro Gly Ser His Tyr Ser Gly Phe Ala
85 90 95

Gly Arg Asp Ala Ser Arg Ala Phe Val Thr Gly Asp Cys Ser Glu Ala
100 105 110

Gly Leu Val Asp Asp Val Ser Asp Leu Ser Ala Ala Glu Met Leu Thr
115 120 125

Leu His Asn Trp Leu Ser Phe Tyr Glu Lys Asn Tyr Val Cys Val Gly
130 135 140

Arg Val Thr Gly Arg Phe Tyr Gly Glu Asp Gly Leu Pro Thr Pro Ala
145 150 155 160

138

Leu Thr Gln Val Glu Ala Ala Ile Thr Arg Gly Leu Glu Ala Asn Lys
 165 170 175
 Leu Gln Leu Gln Glu Lys Gln Thr Phe Pro Pro Cys Asn Ala Glu Trp
 180 185 190
 Ser Ser Ala Arg Gly Ser Arg Leu Trp Cys Ser Gln Lys Ser Gly Gly
 195 200 205
 Val Ser Arg Asp Trp Ile Gly Val Pro Arg Lys Leu Tyr Lys Pro Gly
 210 215 220
 Ala Lys Glu Pro Arg Cys Val Cys Val Arg Thr Thr Gly Pro Pro Ser
 225 230 235 240
 Gly Gln Met Pro Asp Asn Pro Pro His Arg Asn Arg Gly Asp Leu Asp
 245 250 255
 His Pro Asn Leu Ala Glu Tyr Thr Gly Cys Pro Pro Leu Ala Ile Thr
 260 265 270
 Cys Ser Phe Pro Leu
 275

<210> 256
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 256
 Ser Gly Asn Leu Gly Ser Ala Asp Gly Trp Ala Tyr Ile Asp Val Glu
 1 5 10 15
 Val Arg Arg Pro Trp Ala Phe Val Gly Pro Gly Cys Ser Arg Ser Ser
 20 25 30
 Gly Asn Gly Ser
 35

<210> 257
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 257
 Thr Ala Tyr Gly Leu Val Gly Ser Pro Arg Trp Leu Ser Pro Phe His
 1 5 10 15
 Thr Gly Gly Ala Val Ser Leu Pro Arg Arg Pro Arg Gly Pro Gly Pro
 20 25 30
 Val Leu Gly Val
 35

<210> 258
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 258
 Ala Arg Pro Cys Leu Arg Cys Val Leu Arg Pro Glu His Tyr Glu Pro

139

1 5 10 15
Gly Ser His Tyr Ser Gly Phe Ala Gly Arg Asp Ala Ser Arg Ala Phe
 20 25 30

Val Thr Gly Asp
 35

<210> 259
<211> 36
<212> PRT
<213> Homo sapiens

<400> 259
Cys Ser Glu Ala Gly Leu Val Asp Asp Val Ser Asp Leu Ser Ala Ala
1 5 10 15
Glu Met Leu Thr Leu His Asn Trp Leu Ser Phe Tyr Glu Lys Asn Tyr
 20 25 30

Val Cys Val Gly
 35

<210> 260
<211> 36
<212> PRT
<213> Homo sapiens

<400> 260
Arg Val Thr Gly Arg Phe Tyr Gly Glu Asp Gly Leu Pro Thr Pro Ala
1 5 10 15
Leu Thr Gln Val Glu Ala Ala Ile Thr Arg Gly Leu Glu Ala Asn Lys
 20 25 30

Leu Gln Leu Gln
 35

<210> 261
<211> 36
<212> PRT
<213> Homo sapiens

<400> 261
Glu Lys Gln Thr Phe Pro Pro Cys Asn Ala Glu Trp Ser Ser Ala Arg
1 5 10 15
Gly Ser Arg Leu Trp Cys Ser Gln Lys Ser Gly Gly Val Ser Arg Asp
 20 25 30

Trp Ile Gly Val
 35

<210> 262
<211> 29
<212> PRT
<213> Homo sapiens

<400> 262
Pro Arg Lys Leu Tyr Lys Pro Gly Ala Lys Glu Pro Arg Cys Val Cys

140

1 5 10 15
 Val Arg Thr Thr Gly Pro Pro Ser Gly Gln Met Pro Asp
 20 25

<210> 263
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 263
 Asn Pro Pro His Arg Asn Arg Gly Asp Leu Asp His Pro Asn Leu Ala
 1 5 10 15
 Glu Tyr Thr Gly Cys Pro Pro Leu Ala Ile Thr Cys Ser Phe Pro Leu
 20 25 30

<210> 264
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 264
 Arg Asp Asn Asp Tyr Leu Leu His Gly His Arg Pro Pro Met Phe
 1 5 10 15

<210> 265
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 265
 Ser Phe Arg Ala Cys Phe Lys Ser Ile Phe Arg Ile His Thr Glu Thr
 1 5 10 15
 Gly Asn Ile Trp Thr His Leu Leu
 20

<210> 266
 <211> 29
 <212> PRT
 <213> Homo sapiens

<400> 266
 Gly Phe Val Leu Phe Leu Phe Leu Gly Ile Leu Thr Met Leu Arg Pro
 1 5 10 15
 Asn Met Tyr Phe Met Ala Pro Leu Gln Glu Lys Val Val
 20 25

<210> 267
 <211> 457
 <212> PRT
 <213> Homo sapiens

<400> 267

141

Thr Gly Pro Glu Phe Pro Gly Ser Asn Ser Thr Val Ala Arg Arg Ile
 1 5 10 15
 Lys Asp Leu Ala Ala Asp Ile Glu Glu Glu Leu Val Cys Arg Leu Lys
 20 25 30
 Ile Cys Asp Gly Phe Ser Leu Gln Leu Asp Glu Ser Ala Asp Val Ser
 35 40 45
 Gly Leu Ala Val Leu Leu Val Phe Val Arg Tyr Arg Phe Asn Lys Ser
 50 55 60
 Ile Glu Glu Asp Leu Leu Leu Cys Glu Ser Leu Gln Ser Asn Ala Thr
 65 70 75 80
 Gly Glu Glu Ile Phe Asn Cys Ile Asn Ser Phe Met Gln Lys His Glu
 85 90 95
 Ile Glu Trp Glu Lys Cys Val Asp Val Cys Ser Asp Ala Ser Arg Ala
 100 105 110
 Val Asp Gly Lys Ile Ala Glu Ala Val Thr Leu Ile Lys Tyr Val Ala
 115 120 125
 Pro Glu Ser Thr Ser Ser His Cys Leu Leu Tyr Arg His Ala Leu Ala
 130 135 140
 Val Lys Ile Met Pro Thr Ser Leu Lys Asn Val Leu Asp Gln Ala Val
 145 150 155 160
 Gln Ile Ile Asn Tyr Ile Lys Ala Arg Pro His Gln Ser Arg Leu Leu
 165 170 175
 Lys Ile Leu Cys Glu Glu Met Gly Ala Gln His Thr Ala Leu Leu Leu
 180 185 190
 Asn Thr Glu Val Arg Trp Leu Ser Arg Gly Lys Val Leu Val Arg Leu
 195 200 205
 Phe Glu Leu Arg Arg Glu Leu Leu Val Phe Met Asp Ser Ala Phe Arg
 210 215 220
 Leu Ser Asp Cys Leu Thr Asn Ser Ser Trp Leu Leu Arg Leu Ala Tyr
 225 230 235 240
 Leu Ala Asp Ile Phe Thr Lys Leu Asn Glu Val Asn Leu Ser Met Gln
 245 250 255
 Gly Lys Asn Val Thr Val Phe Thr Val Phe Asp Lys Met Ser Ser Leu
 260 265 270
 Leu Arg Lys Leu Glu Phe Trp Ala Ser Ser Val Glu Glu Glu Asn Phe
 275 280 285
 Asp Cys Phe Pro Thr Leu Ser Asp Phe Leu Thr Glu Ile Asn Ser Thr
 290 295 300
 Val Asp Lys Asp Ile Cys Ser Ala Ile Val Gln His Leu Arg Gly Leu
 305 310 315 320
 Arg Ala Thr Leu Leu Lys Tyr Phe Pro Val Thr Asn Asp Asn Asn Ala
 325 330 335
 Trp Val Arg Asn Pro Phe Thr Val Thr Val Lys Pro Ala Ser Leu Val

<400> 270
Leu Asp Gln Ala Val Gln Ile Ile Asn Tyr Ile Lys Ala Arg Pro His
1 5 10 15
Gln Ser Arg Leu Leu Lys Ile Leu Cys Glu Glu Met Gly Ala Gln His
20 25 30

143

Thr Ala Leu Leu
35

<210> 271
<211> 49
<212> PRT
<213> Homo sapiens

<400> 271
Ser Ala Phe Arg Leu Ser Asp Cys Leu Thr Asn Ser Ser Trp Leu Leu
1 5 10 15
Arg Leu Ala Tyr Leu Ala Asp Ile Phe Thr Lys Leu Asn Glu Val Asn
20 25 30
Leu Ser Met Gln Gly Lys Asn Val Thr Val Phe Thr Val Phe Asp Lys
35 40 45

Met

<210> 272
<211> 32
<212> PRT
<213> Homo sapiens

<400> 272
Ser Asp Phe Leu Thr Glu Ile Asn Ser Thr Val Asp Lys Asp Ile Cys
1 5 10 15
Ser Ala Ile Val Gln His Leu Arg Gly Leu Arg Ala Thr Leu Leu Lys
20 25 30

<210> 273
<211> 38
<212> PRT
<213> Homo sapiens

<400> 273
Ser Asp Ser Gln Val Lys Gln Asn Phe Ser Glu Leu Ser Leu Asn Asp
1 5 10 15
Phe Trp Ser Ser Leu Ile Gln Glu Tyr Pro Ser Ile Ala Arg Arg Ala
20 25 30
Val Arg Val Leu Leu Pro
35

<210> 274
<211> 26
<212> PRT
<213> Homo sapiens

<400> 274
Asp Ser Arg Ile Ser Leu Leu Val Asn Asn Ala Gly Val Gly Ala Thr
1 5 10 15

144

Ala Ser Leu Leu Glu Ser Asp Ala Asp Lys
 20 25

<210> 275
 <211> 146
 <212> PRT
 <213> Homo sapiens

<400> 275
 Gly Thr Pro Ala Gly Thr Gly Pro Glu Phe Pro Gly Arg Pro Thr Arg
 1 5 10 15
 Pro Ser Arg Thr Glu Ser Ala Gln Thr Thr Gln His Ser Pro Leu Arg
 20 25 30
 Pro Leu Trp Arg Leu Lys Arg Asp Ser Ser Pro Cys His Pro Gln Thr
 35 40 45
 Arg Ala Asp Trp Gly Val Cys Pro Pro Trp Gly Gly Ala Ala Gln Gly
 50 55 60
 Leu Arg Pro Gly Cys His Leu Ala Pro Arg Arg Cys Leu Cys Pro Gly
 65 70 75 80
 Ser Cys Cys Pro Trp His Trp Ala Glu Ala Gln Trp Ser Phe Leu Trp
 85 90 95
 Arg Gly Leu Trp Gly Leu Arg Thr Leu Pro Thr Ala Leu Arg Ala Ser
 100 105 110
 Pro Ala Ala Ser Gly Thr Val Thr Tyr Ser Ala Cys Leu Gly Thr Ser
 115 120 125
 Cys Leu Leu Arg Ala Pro Cys Trp Arg Leu Arg Thr Cys Arg Gln Ser
 130 135 140
 Trp Cys
 145

<210> 276
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 276
 Gly Thr Pro Ala Gly Thr Gly Pro Glu Phe Pro Gly Arg Pro Thr Arg
 1 5 10 15
 Pro Ser Arg Thr Glu Ser Ala Gln Thr Thr Gln His
 20 25

<210> 277
 <211> 30
 <212> PRT
 <213> Homo sapiens

<400> 277
 Ser Pro Leu Arg Pro Leu Trp Arg Leu Lys Arg Asp Ser Ser Pro Cys
 1 5 10 15

145

His Pro Gln Thr Arg Ala Asp Trp Gly Val Cys Pro Pro Trp
20 25 30

<210> 278
<211> 30
<212> PRT
<213> Homo sapiens

<400> 278
Gly Gly Ala Ala Gln Gly Leu Arg Pro Gly Cys His Leu Ala Pro Arg
1 5 10 15

Arg Cys Leu Cys Pro Gly Ser Cys Cys Pro Trp His Trp Ala
20 25 30

<210> 279
<211> 30
<212> PRT
<213> Homo sapiens

<400> 279
Glu Ala Gln Trp Ser Phe Leu Trp Arg Gly Leu Trp Gly Leu Arg Thr
1 5 10 15

Leu Pro Thr Ala Leu Arg Ala Ser Pro Ala Ala Ser Gly Thr
20 25 30

<210> 280
<211> 28
<212> PRT
<213> Homo sapiens

<400> 280
Val Thr Tyr Ser Ala Cys Leu Gly Thr Ser Cys Leu Leu Arg Ala Pro
1 5 10 15

Cys Trp Arg Leu Arg Thr Cys Arg Gln Ser Trp Cys
20 25

<210> 281
<211> 11
<212> PRT
<213> Homo sapiens

<400> 281
Pro Pro Arg Pro Ser Thr Ser Gly Gln Trp Gly
1 5 10

<210> 282
<211> 11
<212> PRT
<213> Homo sapiens

<400> 282
Arg Arg Ser Pro Phe Thr Ser Ala Gln Thr Gly
1 5 10

<210> 283

146

<211> 23
 <212> PRT
 <213> Homo sapiens

<400> 283
 Gly Thr Gly Trp Asp Phe Gly Leu Ala Ala Val Cys Leu Arg Ala Ala
 1 5 10 15
 Glu Val Ala Gly Ser Phe Lys
 20

<210> 284
 <211> 146
 <212> PRT
 <213> Homo sapiens

<400> 284
 Gly Tyr Arg Arg Val Phe Glu Glu Tyr Met Arg Val Ile Ser Gln Arg
 1 5 10 15
 Tyr Pro Asp Ile Arg Ile Glu Gly Glu Asn Tyr Leu Pro Gln Pro Ile
 20 25 30
 Tyr Arg His Ile Ala Ser Phe Leu Ser Val Phe Lys Leu Val Leu Ile
 35 40 45
 Gly Leu Ile Ile Val Gly Lys Asp Pro Phe Ala Phe Phe Gly Met Gln
 50 55 60
 Ala Pro Ser Ile Trp Gln Trp Gly Gln Glu Asn Lys Val Tyr Ala Cys
 65 70 75 80
 Met Met Val Phe Phe Leu Ser Asn Met Ile Glu Asn Gln Cys Met Ser
 85 90 95
 Thr Gly Ala Phe Glu Ile Thr Leu Asn Asp Val Pro Val Trp Ser Lys
 100 105 110
 Leu Glu Ser Gly His Leu Pro Ser Met Gln Gln Leu Val Gln Ile Leu
 115 120 125
 Asp Asn Glu Met Lys Leu Asn Val His Met Asp Ser Ile Pro His His
 130 135 140
 Arg Ser
 145

<210> 285
 <211> 34
 <212> PRT
 <213> Homo sapiens

<400> 285
 Gly Tyr Arg Arg Val Phe Glu Glu Tyr Met Arg Val Ile Ser Gln Arg
 1 5 10 15
 Tyr Pro Asp Ile Arg Ile Glu Gly Glu Asn Tyr Leu Pro Gln Pro Ile
 20 25 30
 Tyr Arg

147

<210> 286
<211> 34
<212> PRT
<213> Homo sapiens

<400> 286
His Ile Ala Ser Phe Leu Ser Val Phe Lys Leu Val Leu Ile Gly Leu
1 5 10 15
Ile Ile Val Gly Lys Asp Pro Phe Ala Phe Phe Gly Met Gln Ala Pro
20 25 30

Ser Ile

<210> 287
<211> 34
<212> PRT
<213> Homo sapiens

<400> 287
Trp Gln Trp Gly Gln Glu Asn Lys Val Tyr Ala Cys Met Met Val Phe
1 5 10 15
Phe Leu Ser Asn Met Ile Glu Asn Gln Cys Met Ser Thr Gly Ala Phe
20 25 30

Glu Ile

<210> 288
<211> 36
<212> PRT
<213> Homo sapiens

<400> 288
Thr Leu Asn Asp Val Pro Val Trp Ser Lys Leu Glu Ser Gly His Leu
1 5 10 15
Pro Ser Met Gln Gln Leu Val Gln Ile Leu Asp Asn Glu Met Lys Leu
20 25 30

Asn Val His Met
35

<210> 289
<211> 8
<212> PRT
<213> Homo sapiens

<400> 289
Asp Ser Ile Pro His His Arg Ser
1 5

<210> 290
<211> 30
<212> PRT
<213> Homo sapiens

148

<400> 290

Gly Arg Ala Arg Gly Arg Pro Pro Gly Pro Glu Ala Ala Pro Ala Ser
 1 5 10 15

Leu Ser Val Ser Leu Arg Arg Glu Val His Ser Arg Gly Glu
 20 25 30

<210> 291

<211> 389

<212> PRT

<213> Homo sapiens

<400> 291

Met Ile Ser Leu Pro Gly Pro Leu Val Thr Asn Leu Leu Arg Phe Leu
 1 5 10 15

Phe Leu Gly Leu Ser Ala Leu Ala Pro Pro Ser Arg Ala Gln Leu Gln
 20 25 30

Leu His Leu Pro Ala Asn Arg Leu Gln Ala Val Glu Gly Gly Glu Val
 35 40 45

Val Leu Pro Ala Trp Tyr Thr Leu His Gly Glu Val Ser Ser Ser Gln
 50 55 60

Pro Trp Glu Val Pro Phe Val Met Trp Phe Phe Lys Gln Lys Glu Lys
 65 70 75 80

Glu Asp Gln Val Leu Ser Tyr Ile Asn Gly Val Thr Thr Ser Lys Pro
 85 90 95

Gly Val Ser Leu Val Tyr Ser Met Pro Ser Arg Asn Leu Ser Leu Arg
 100 105 110

Leu Glu Gly Leu Gln Glu Lys Asp Ser Gly Pro Tyr Ser Cys Ser Val
 115 120 125

Asn Val Gln Asn Lys Gln Gly Lys Ser Arg Gly His Ser Ile Lys Thr
 130 135 140

Leu Glu Leu Asn Val Leu Val Pro Pro Ala Pro Pro Ser Cys Arg Leu
 145 150 155 160

Gln Gly Val Pro His Val Gly Ala Asn Val Thr Leu Ser Cys Gln Ser
 165 170 175

Pro Arg Ser Lys Pro Ala Val Gln Tyr Gln Trp Asp Arg Gln Leu Pro
 180 185 190

Ser Phe Gln Thr Phe Phe Ala Pro Ala Leu Asp Val Ile Arg Gly Ser
 195 200 205

Leu Ser Leu Thr Asn Leu Ser Ser Ser Met Ala Gly Val Tyr Val Cys
 210 215 220

Lys Ala His Asn Glu Val Gly Thr Ala Gln Cys Asn Val Thr Leu Glu
 225 230 235 240

Val Ser Thr Gly Pro Gly Ala Ala Val Val Ala Gly Ala Val Val Gly
 245 250 255

Thr Leu Val Gly Leu Gly Leu Leu Ala Gly Leu Val Leu Leu Tyr His
 260 265 270

149

Arg Arg Gly Lys Ala Leu Glu Glu Pro Ala Asn Asp Ile Lys Glu Asp
 275 280 285

Ala Ile Ala Pro Arg Thr Leu Pro Trp Pro Lys Ser Ser Asp Thr Ile
 290 295 300

Ser Lys Asn Gly Thr Leu Ser Ser Val Thr Ser Ala Arg Ala Leu Arg
 305 310 315 320

Pro Pro His Gly Pro Pro Arg Pro Gly Ala Leu Thr Pro Thr Pro Ser
 325 330 335

Leu Ser Ser Gln Ala Leu Pro Ser Pro Arg Leu Pro Thr Thr Asp Gly
 340 345 350

Ala His Pro Gln Pro Ile Ser Pro Ile Pro Gly Gly Val Ser Ser Ser
 355 360 365

Gly Leu Ser Arg Met Gly Ala Val Pro Val Met Val Pro Ala Gln Ser
 370 375 380

Gln Ala Gly Ser Leu
 385

<210> 292
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 292
 Met Ile Ser Leu Pro Gly Pro Leu Val Thr Asn Leu Leu Arg Phe Leu
 1 5 10 15

Phe Leu Gly Leu Ser Ala Leu Ala Pro Pro Ser Arg Ala Gln Leu Gln
 20 25 30

Leu His Leu
 35

<210> 293
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 293
 Pro Ala Asn Arg Leu Gln Ala Val Glu Gly Gly Glu Val Val Leu Pro
 1 5 10 15

Ala Trp Tyr Thr Leu His Gly Glu Val Ser Ser Ser Gln Pro Trp Glu
 20 25 30

Val Pro Phe
 35

<210> 294
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 294

150

Val Met Trp Phe Phe Lys Gln Lys Glu Lys Glu Asp Gln Val Leu Ser
1 5 10 15

Tyr Ile Asn Gly Val Thr Thr Ser Lys Pro Gly Val Ser Leu Val Tyr
20 25 30

Ser Met Pro
35

<210> 295
<211> 35
<212> PRT
<213> Homo sapiens

<400> 295
Ser Arg Asn Leu Ser Leu Arg Leu Glu Gly Leu Gln Glu Lys Asp Ser
1 5 10 15

Gly Pro Tyr Ser Cys Ser Val Asn Val Gln Asn Lys Gln Gly Lys Ser
20 25 30

Arg Gly His
35

<210> 296
<211> 35
<212> PRT
<213> Homo sapiens

<400> 296
Ser Ile Lys Thr Leu Glu Leu Asn Val Leu Val Pro Pro Ala Pro Pro
1 5 10 15

Ser Cys Arg Leu Gln Gly Val Pro His Val Gly Ala Asn Val Thr Leu
20 25 30

Ser Cys Gln
35

<210> 297
<211> 35
<212> PRT
<213> Homo sapiens

<400> 297
Ser Pro Arg Ser Lys Pro Ala Val Gln Tyr Gln Trp Asp Arg Gln Leu
1 5 10 15

Pro Ser Phe Gln Thr Phe Phe Ala Pro Ala Leu Asp Val Ile Arg Gly
20 25 30

Ser Leu Ser
35

<210> 298
<211> 35
<212> PRT
<213> Homo sapiens

<400> 298

151

Leu Thr Asn Leu Ser Ser Ser Met Ala Gly Val Tyr Val Cys Lys Ala
1 5 10 15

His Asn Glu Val Gly Thr Ala Gln Cys Asn Val Thr Leu Glu Val Ser
20 25 30

Thr Gly Pro
35

<210> 299
<211> 35
<212> PRT
<213> Homo sapiens

<400> 299
Gly Ala Ala Val Val Ala Gly Ala Val Val Gly Thr Leu Val Gly Leu
1 5 10 15

Gly Leu Leu Ala Gly Leu Val Leu Leu Tyr His Arg Arg Gly Lys Ala
20 25 30

Leu Glu Glu
35

<210> 300
<211> 35
<212> PRT
<213> Homo sapiens

<400> 300
Pro Ala Asn Asp Ile Lys Glu Asp Ala Ile Ala Pro Arg Thr Leu Pro
1 5 10 15

Trp Pro Lys Ser Ser Asp Thr Ile Ser Lys Asn Gly Thr Leu Ser Ser
20 25 30

Val Thr Ser
35

<210> 301
<211> 35
<212> PRT
<213> Homo sapiens

<400> 301
Ala Arg Ala Leu Arg Pro Pro His Gly Pro Pro Arg Pro Gly Ala Leu
1 5 10 15

Thr Pro Thr Pro Ser Leu Ser Ser Gln Ala Leu Pro Ser Pro Arg Leu
20 25 30

Pro Thr Thr
35

<210> 302
<211> 39
<212> PRT
<213> Homo sapiens

<400> 302

152

Asp Gly Ala His Pro Gln Pro Ile Ser Pro Ile Pro Gly Gly Val Ser
 1 5 10 15

Ser Ser Gly Leu Ser Arg Met Gly Ala Val Pro Val Met Val Pro Ala
 20 25 30

Gln Ser Gln Ala Gly Ser Leu
 35

<210> 303
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 303
 Gly Ser Ser Phe Val Val Ser Glu Gly Ser Tyr Leu Asp Ile Ser Asp
 1 5 10 15

Trp Leu Asn Pro Ala Lys Leu Ser Leu Tyr Tyr
 20 25

<210> 304
 <211> 12
 <212> PRT
 <213> Homo sapiens

<400> 304
 Leu Asp Ile Ser Asp Trp Leu Asn Pro Ala Lys Leu
 1 5 10

<210> 305
 <211> 11
 <212> PRT
 <213> Homo sapiens

<400> 305
 Ser Asp Trp Leu Asn Pro Ala Lys Leu Ser Leu
 1 5 10

<210> 306
 <211> 11
 <212> PRT
 <213> Homo sapiens

<400> 306
 Ser Asp Trp Leu Asn Pro Ala Lys Leu Ser Leu
 1 5 10

<210> 307
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 307
 Gly Ser Ser Phe Val Val Ser Glu Gly Ser Tyr Leu Asp Ile Ser Asp
 1 5 10 15

Trp Leu Asn Pro Ala Lys Leu Ser Leu Tyr Tyr
 20 25

<210> 308
<211> 12
<212> PRT
<213> Homo sapiens

<400> 308
Leu Asp Ile Ser Asp Trp Leu Asn Pro Ala Lys Leu
1 5 10

<210> 309
<211> 11
<212> PRT
<213> Homo sapiens

<400> 309
Ser Asp Trp Leu Asn Pro Ala Lys Leu Ser Leu
1 5 10

<210> 310
<211> 13
<212> PRT
<213> Homo sapiens

<400> 310
Asp Ala Cys Glu Gln Leu Cys Asp Pro Glu Thr Gly Glu
1 5 10

<210> 311
<211> 21
<212> PRT
<213> Homo sapiens

<400> 311
Glu Gly Lys Ile Lys Ile Cys Glu Lys Lys Ala Ile Lys Val Ile Leu
1 5 10 15

His Thr Cys Asn Ser
20

<210> 312
<211> 23
<212> PRT
<213> Homo sapiens

<400> 312
Asn Ser Ala Arg Val Glu Phe Phe Ile Pro Pro Leu Arg Ile Thr Gln
1 5 10 15

Lys Val Arg Ser Thr Lys Ser
20

<210> 313
<211> 123
<212> PRT
<213> Homo sapiens

<400> 313

154

Met Met Val Trp Asn Leu Phe Pro Cys Phe Pro Pro Leu Leu Leu Leu
 1 5 10 15
 Gln Phe Ile Asp Cys Gln Gln Ser Ser Glu Ile Glu Gln Gly Phe Thr
 20 25 30
 Arg Ser Leu Leu Gly His Pro Ile Phe Phe Cys Pro Asp Pro Cys Trp
 35 40 45
 Gln Ser Cys Met Asn Cys Val Ile Leu Ser Val Leu Ser Phe Phe Phe
 50 55 60
 Leu Ile Arg Trp Ile Ser Lys Ile Val Ala Val Gln Lys Leu Glu Ser
 65 70 75 80
 Ser Ser Arg Arg Lys Pro Ile Leu Phe Leu Ile Ile Ser Cys Glu Ile
 85 90 95
 Ala Ser Phe Ile His Leu Phe Leu Ser Gln Met Ser Ala Glu Cys Cys
 100 105 110
 Cys Phe Tyr Leu Val Ile Leu Ile Cys Lys Tyr
 115 120

<210> 314
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 314
 Met Met Val Trp Asn Leu Phe Pro Cys Phe Pro Pro Leu Leu Leu Leu
 1 5 10 15
 Gln Phe Ile Asp Cys Gln Gln Ser Ser Glu Ile Glu
 20 25

<210> 315
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 315
 Gln Gly Phe Thr Arg Ser Leu Leu Gly His Pro Ile Phe Phe Cys Pro
 1 5 10 15
 Asp Pro Cys Trp Gln Ser Cys Met Asn Cys Val Ile
 20 25

<210> 316
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 316
 Leu Ser Val Leu Ser Phe Phe Phe Leu Ile Arg Trp Ile Ser Lys Ile
 1 5 10 15
 Val Ala Val Gln Lys Leu Glu Ser Ser Ser Arg Arg Lys Pro Ile Leu
 20 25 30
 Phe Leu Ile

155

35

<210> 317
<211> 32
<212> PRT
<213> Homo sapiens

<400> 317
Ile Ser Cys Glu Ile Ala Ser Phe Ile His Leu Phe Leu Ser Gln Met
1 5 10 15
Ser Ala Glu Cys Cys Cys Phe Tyr Leu Val Ile Leu Ile Cys Lys Tyr
20 25 30

<210> 318
<211> 59
<212> PRT
<213> Homo sapiens

<400> 318
Lys Val Asp Thr Pro Arg Arg His Phe Cys Pro Glu Ile Ser Phe Phe
1 5 10 15
Leu Thr Pro Leu Pro Gln Ser Ala Arg Asn Ser Thr Val Arg Asn Ala
20 25 30
Leu Ser Gly Leu Lys Asn Leu Thr Pro Ala Met Ile Ser Thr Val Ser
35 40 45
Lys Gln Asp Thr Ser Lys Leu Gly Glu Glu Glu
50 55

<210> 319
<211> 13
<212> PRT
<213> Homo sapiens

<400> 319
Leu Leu Leu Cys Pro Trp Trp Leu Cys Phe Asp Trp Ser
1 5 10

<210> 320
<211> 270
<212> PRT
<213> Homo sapiens

<400> 320
Met Gly Cys Ile Pro Leu Ile Lys Ser Ile Ser Asp Trp Arg Val Ile
1 5 10 15
Ala Leu Ala Ala Leu Trp Phe Cys Leu Ile Gly Leu Ile Cys Gln Ala
20 25 30
Leu Cys Ser Glu Asp Gly His Lys Arg Arg Ile Leu Thr Leu Gly Leu
35 40 45
Gly Phe Leu Val Ile Pro Phe Leu Pro Ala Ser Asn Leu Phe Phe Arg

156

50 55 60
 Val Gly Phe Val Val Ala Glu Cys Val Leu Tyr Leu Pro Ser Ile Gly
 65 70 75 80
 Tyr Cys Val Leu Leu Thr Phe Gly Phe Gly Ala Leu Ser Lys His Thr
 85 90 95
 Lys Lys Lys Lys Leu Ile Ala Ala Val Val Leu Gly Ile Leu Phe Ile
 100 105 110
 Asn Thr Leu Arg Cys Val Leu Arg Thr Ala Lys Trp Arg Ser Glu Glu
 115 120 125
 Gln Leu Phe Arg Ser Ala Leu Ser Val Cys Pro Leu Asn Ala Lys Val
 130 135 140
 His Tyr Asn Ile Gly Lys Asn Leu Ala Asp Lys Gly Asn Gln Thr Ala
 145 150 155 160
 Ala Ile Arg Tyr Tyr Arg Glu Ala Val Arg Leu Asn Pro Lys Tyr Val
 165 170 175
 His Ala Met Asn Asn Leu Gly Asn Ile Leu Lys Glu Arg Asn Glu Leu
 180 185 190
 Gln Glu Ala Glu Glu Leu Leu Ser Leu Ala Val Gln Ile Gln Pro Asp
 195 200 205
 Phe Ala Ala Ala Trp Met Asn Leu Gly Ile Val Gln Asn Ser Leu Lys
 210 215 220
 Arg Phe Glu Thr Ala Glu Gln Asn Tyr Arg Thr Ala Ile Lys His Arg
 225 230 235 240
 Arg Lys Tyr Pro Asp Cys Tyr Tyr Asn Leu Gly Arg Leu Val Arg Thr
 245 250 255
 Gly Cys Pro Val Pro Val Glu Gly Lys Met Gly Tyr Phe Ser
 260 265 270

<210> 321
 <211> 38
 <212> PRT
 <213> Homo sapiens

<400> 321
 Met Gly Cys Ile Pro Leu Ile Lys Ser Ile Ser Asp Trp Arg Val Ile
 1 5 10 15

Ala Leu Ala Ala Leu Trp Phe Cys Leu Ile Gly Leu Ile Cys Gln Ala
 20 25 30

Leu Cys Ser Glu Asp Gly
 35

<210> 322
 <211> 38
 <212> PRT
 <213> Homo sapiens

<400> 322

157

His Lys Arg Arg Ile Leu Thr Leu Gly Leu Gly Phe Leu Val Ile Pro
1 5 10 15

Phe Leu Pro Ala Ser Asn Leu Phe Phe Arg Val Gly Phe Val Val Ala
20 25 30

Glu Cys Val Leu Tyr Leu
35

<210> 323

<211> 38

<212> PRT

<213> Homo sapiens

<400> 323

Pro Ser Ile Gly Tyr Cys Val Leu Leu Thr Phe Gly Phe Gly Ala Leu
1 5 10 15

Ser Lys His Thr Lys Lys Lys Lys Leu Ile Ala Ala Val Val Leu Gly
20 25 30

Ile Leu Phe Ile Asn Thr
35

<210> 324

<211> 38

<212> PRT

<213> Homo sapiens

<400> 324

Pro Ser Ile Gly Tyr Cys Val Leu Leu Thr Phe Gly Phe Gly Ala Leu
1 5 10 15

Ser Lys His Thr Lys Lys Lys Lys Leu Ile Ala Ala Val Val Leu Gly
20 25 30

Ile Leu Phe Ile Asn Thr
35

<210> 325

<211> 38

<212> PRT

<213> Homo sapiens

<400> 325

Leu Arg Cys Val Leu Arg Thr Ala Lys Trp Arg Ser Glu Glu Gln Leu
1 5 10 15

Phe Arg Ser Ala Leu Ser Val Cys Pro Leu Asn Ala Lys Val His Tyr
20 25 30

Asn Ile Gly Lys Asn Leu
35

<210> 326

<211> 38

<212> PRT

<213> Homo sapiens

<400> 326

Ala Asp Lys Gly Asn Gln Thr Ala Ala Ile Arg Tyr Tyr Arg Glu Ala
1 5 10 15

Val Arg Leu Asn Pro Lys Tyr Val His Ala Met Asn Asn Leu Gly Asn
20 25 30

Ile Leu Lys Glu Arg Asn
35

<210> 327
<211> 38
<212> PRT
<213> Homo sapiens

<400> 327
Glu Leu Gln Glu Ala Glu Glu Leu Leu Ser Leu Ala Val Gln Ile Gln
1 5 10 15

Pro Asp Phe Ala Ala Ala Trp Met Asn Leu Gly Ile Val Gln Asn Ser
20 25 30

Leu Lys Arg Phe Glu Thr
35

<210> 328
<211> 42
<212> PRT
<213> Homo sapiens

<400> 328
Ala Glu Gln Asn Tyr Arg Thr Ala Ile Lys His Arg Arg Lys Tyr Pro
1 5 10 15

Asp Cys Tyr Tyr Asn Leu Gly Arg Leu Val Arg Thr Gly Cys Pro Val
20 25 30

Pro Val Glu Gly Lys Met Gly Tyr Phe Ser
35 40

<210> 329
<211> 26
<212> PRT
<213> Homo sapiens

<400> 329
Pro Thr Arg Pro Pro Thr Arg Pro Leu Ser Phe Thr Phe Thr Lys Gln
1 5 10 15

Thr Ser Ser Thr Cys Leu Ser Leu His Phe
20 25

<210> 330
<211> 50
<212> PRT
<213> Homo sapiens

<400> 330
Leu Glu Cys Val Leu Leu Ile Cys Phe Arg Ala Met Ser Ala Ile Tyr
1 5 10 15

159

Thr His Thr Ser Ile Gly Asn Ala Gln Lys Leu Phe Thr Asp Gly Ser
 20 25 30

Ala Phe Arg Arg Val Arg Glu Pro Leu Pro Lys Glu Gly Lys Ser Trp
 35 40 45

Pro Gln
 50

<210> 331
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 331
 Lys Gln Asn Leu Thr Asn Leu Asp Val Pro Val Gln Tyr His Val Ala
 1 5 10 15

Leu Ser Asp Lys Val Lys
 20

<210> 332
 <211> 117
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (71)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 332
 Pro Ser Cys Pro Pro Glu Met Lys Lys Glu Leu Pro Val Asp Ser Cys
 1 5 10 15

Leu Pro Arg Ser Leu Glu Leu His Pro Gln Lys Met Asp Pro Lys Arg
 20 25 30

Gln His Ile Gln Leu Leu Ser Ser Leu Thr Glu Cys Leu Thr Val Asp
 35 40 45

Pro Leu Ser Ala Ser Val Trp Arg Gln Leu Tyr Pro Lys His Leu Ser
 50 55 60

Gln Ser Ser Leu Leu Leu Xaa His Leu Leu Ser Ser Trp Glu Gln Ile
 65 70 75 80

Pro Lys Lys Val Gln Lys Ser Leu Gln Glu Thr Ile Gln Ser Leu Lys
 85 90 95

Leu Thr Asn Gln Glu Leu Leu Arg Lys Gly Ser Ser Asn Asn Gln Asp
 100 105 110

Val Val Thr Cys Asp
 115

<210> 333
 <211> 103
 <212> PRT
 <213> Homo sapiens

160

<400> 333

Lys Ala Pro Tyr Ser Trp Leu Ala Asp Ser Trp Pro His Pro Ser Arg
 1 5 10 15
 Ser Pro Ser Ala Gln Glu Pro Arg Gly Ser Cys Cys Pro Ser Asn Pro
 20 25 30
 Asp Pro Asp Asp Arg Tyr Tyr Asn Glu Ala Gly Ile Ser Leu Tyr Leu
 35 40 45
 Ala Gln Thr Ala Arg Gly Thr Ala Ala Pro Gly Glu Gly Pro Val Tyr
 50 55 60
 Ser Thr Ile Asp Pro Ala Gly Glu Glu Leu Gln Thr Phe His Gly Gly
 65 70 75 80
 Phe Pro Gln His Pro Ser Gly Asp Leu Gly Pro Trp Ser Gln Tyr Ala
 85 90 95
 Pro Pro Glu Trp Ser Gln Gly
 100

<210> 334

<211> 43

<212> PRT

<213> Homo sapiens

<400> 334

Leu Gln Gln Thr Met Gln Ala Met Leu His Phe Gly Gly Arg Leu Ala
 1 5 10 15
 Gln Ser Leu Arg Gly Thr Ser Lys Glu Ala Ala Ser Asp Pro Ser Asp
 20 25 30
 Ser Pro Asn Leu Pro Thr Pro Gly Ser Trp Trp
 35 40

<210> 335

<211> 45

<212> PRT

<213> Homo sapiens

<400> 335

Glu Gln Leu Thr Gln Ala Ser Arg Val Tyr Ala Ser Gly Gly Thr Glu
 1 5 10 15
 Gly Phe Pro Leu Ser Arg Trp Ala Pro Gly Arg His Gly Thr Ala Ala
 20 25 30
 Glu Glu Gly Ala Gln Glu Arg Pro Leu Pro Thr Asp Glu
 35 40 45

<210> 336

<211> 45

<212> PRT

<213> Homo sapiens

<400> 336

Met Ala Pro Gly Arg Gly Leu Trp Leu Gly Arg Leu Phe Gly Val Pro
 1 5 10 15

Gly Gly Pro Ala Glu Asn Glu Asn Gly Ala Leu Lys Ser Arg Arg Pro
 20 25 30

Ser Ser Trp Leu Pro Pro Thr Val Ser Val Leu Ala Leu
 35 40 45

<210> 337

<211> 44

<212> PRT

<213> Homo sapiens

<400> 337

Val Lys Arg Gly Ala Pro Pro Glu Met Pro Ser Pro Gln Glu Leu Glu
 1 5 10 15

Ala Ser Ala Pro Arg Met Val Gln Thr His Arg Ala Val Arg Ala Leu
 20 25 30

Cys Asp His Thr Ala Ala Arg Pro Asp Gln Leu Ser
 35 40

<210> 338

<211> 38

<212> PRT

<213> Homo sapiens

<400> 338

Phe Arg Arg Gly Glu Val Leu Arg Val Ile Thr Thr Val Asp Glu Asp
 1 5 10 15

Trp Leu Arg Cys Gly Arg Asp Gly Met Glu Gly Leu Val Pro Val Gly
 20 25 30

Tyr Thr Ser Leu Val Leu
 35

<210> 339

<211> 215

<212> PRT

<213> Homo sapiens

<400> 339

Leu Gln Gln Thr Met Gln Ala Met Leu His Phe Gly Gly Arg Leu Ala
 1 5 10 15

Gln Ser Leu Arg Gly Thr Ser Lys Glu Ala Ala Ser Asp Pro Ser Asp
 20 25 30

Ser Pro Asn Leu Pro Thr Pro Gly Ser Trp Trp Glu Gln Leu Thr Gln
 35 40 45

Ala Ser Arg Val Tyr Ala Ser Gly Gly Thr Glu Gly Phe Pro Leu Ser
 50 55 60

Arg Trp Ala Pro Gly Arg His Gly Thr Ala Ala Glu Glu Gly Ala Gln
 65 70 75 80

Glu Arg Pro Leu Pro Thr Asp Glu Met Ala Pro Gly Arg Gly Leu Trp
 85 90 95

Leu Gly Arg Leu Phe Gly Val Pro Gly Gly Pro Ala Glu Asn Glu Asn

```

<400> 341
Phe  Gln  Ser  Val  Tyr  His  Met  Lys  Leu  Gln  Ser  Ser  Asn  Leu  Pro  Ala
      1          5          10          15
Ser  Val  Tyr  Gly  Asn  Asn  Leu  Asn  Cys  Ile  Asn  Ser  Ser  Ser  Ser
      20          25          30

```

163

<210> 342
<211> 37
<212> PRT
<213> Homo sapiens

<400> 342
Phe Asp Phe Ile Ala Ser Leu Leu Lys Ala Asn Arg Leu Ser Leu Gln
1 5 10 15
Thr Cys Glu Leu Leu Leu Ala Ala Ala Leu Leu Pro Ser Glu Arg Tyr
20 25 30

Lys Ala Ile Ser Ile
35

<210> 343
<211> 25
<212> PRT
<213> Homo sapiens

<400> 343
Ile Asp Leu Ser Phe Pro Ser Thr Asn Val Ser Leu Glu Asp Arg Asn
1 5 10 15
Thr Thr Lys Pro Ser Val Asn Val Gly
20 25

<210> 344
<211> 15
<212> PRT
<213> Homo sapiens

<400> 344
Leu Asn Ile Leu Ile Ser Leu Thr Val Ser Ser His Cys Lys Leu
1 5 10 15

<210> 345
<211> 13
<212> PRT
<213> Homo sapiens

<400> 345
Ile Asn Tyr His Ser Gly Phe Ile His Gln Phe Leu Ala
1 5 10

<210> 346
<211> 11
<212> PRT
<213> Homo sapiens

<400> 346
Met Ala Asn Asn Ser Leu Ser Ser Gln Phe Ile
1 5 10

<210> 347
<211> 46
<212> PRT
<213> Homo sapiens

<400> 347

Pro Phe Trp Ala Ala Glu Ser Ala Leu Asp Phe His Trp Pro Phe Gly
 1 5 10 15

Gly Ala Leu Cys Lys Met Val Leu Thr Ala Thr Val Leu Asn Val Tyr
 20 25 30

Ala Ser Ile Phe Leu Ile Thr Ala Leu Ser Val Ala Arg Tyr
 35 40 45

<210> 348

<211> 12

<212> PRT

<213> Homo sapiens

<400> 348

Thr His Ala Asp Lys Asn Gln Val Arg Asn Ser Asn
 1 5 10

<210> 349

<211> 15

<212> PRT

<213> Homo sapiens

<400> 349

Gln Phe Leu Ser Trp Glu Gln Cys Thr Gly Asn Thr Glu Ser Gln
 1 5 10 15

<210> 350

<211> 13

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (9)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 350

Val Arg Arg Pro Lys Ala Lys Gly Xaa Gln Thr Ser Asn
 1 5 10

<210> 351

<211> 19

<212> PRT

<213> Homo sapiens

<400> 351

Pro Thr Gln Leu Asn Lys His Lys Pro Thr Thr Lys Glu Arg Arg Arg
 1 5 10 15

Lys Gly Leu

<210> 352

<211> 9

<212> PRT

<213> Homo sapiens

165

<400> 352

Leu Ile Ser Lys His Glu Asn Ile Tyr
 1 5

<210> 353

<211> 27

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (6)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (8)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (22)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 353

Thr Leu Tyr Ile Xaa Xaa Met Xaa Thr Gln Thr Trp Arg Asp Gln Gly
 1 5 10 15

Arg Cys Gly Arg Asp Xaa Ile Asn Cys Ile Val
 20 25

<210> 354

<211> 33

<212> PRT

<213> Homo sapiens

<400> 354

Ser Leu Cys Thr Pro Gly Arg Gly Trp Glu Glu Ser Trp Gly Ser Ser
 1 5 10 15

Leu Pro Asn Leu Thr Gly Trp Ser Val Ser Ser Leu Asp Asn Asn Asp
 20 25 30

Val

<210> 355

<211> 204

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (107)

<223> Xaa equals any of the naturally occurring L-amino acids

166

<400> 355

Met Gln Val Ala Leu Lys Glu Asp Leu Asp Ala Leu Lys Glu Lys Phe
 1 5 10 15

Arg Thr Met Glu Ser Asn Gln Lys Ser Ser Phe Gln Glu Ile Pro Lys
 20 25 30

Leu Asn Glu Glu Leu Leu Ser Lys Gln Lys Gln Leu Glu Lys Ile Glu
 35 40 45

Ser Gly Glu Met Gly Leu Asn Lys Val Trp Ile Asn Ile Thr Glu Met
 50 55 60

Asn Lys Gln Ile Ser Leu Leu Thr Ser Ala Val Asn His Leu Lys Ala
 65 70 75 80

Asn Val Lys Ser Ala Ala Asp Leu Ile Ser Leu Pro Thr Thr Val Glu
 85 90 95

Gly Leu Gln Lys Ser Val Ala Ser Ile Gly Xaa Thr Leu Asn Ser Val
 100 105 110

His Leu Ala Val Glu Ala Leu Gln Lys Thr Val Asp Glu His Lys Lys
 115 120 125

Thr Met Glu Leu Leu Gln Ser Asp Met Asn Gln His Phe Leu Lys Glu
 130 135 140

Thr Pro Gly Ser Asn Gln Ile Ile Pro Ser Pro Ser Ala Thr Ser Glu
 145 150 155 160

Leu Asp Asn Lys Thr His Ser Glu Asn Leu Lys Gln Met Gly Asp Arg
 165 170 175

Ser Ala Thr Leu Lys Arg Gln Ser Leu Asp Gln Val Thr Asn Arg Thr
 180 185 190

Asp Thr Val Lys Ile Gln Ser Ile Lys Lys Glu Gly
 195 200

<210> 356

<211> 43

<212> PRT

<213> Homo sapiens

<400> 356

Met Gln Val Ala Leu Lys Glu Asp Leu Asp Ala Leu Lys Glu Lys Phe
 1 5 10 15

Arg Thr Met Glu Ser Asn Gln Lys Ser Ser Phe Gln Glu Ile Pro Lys
 20 25 30

Leu Asn Glu Glu Leu Leu Ser Lys Gln Lys Gln
 35 40

<210> 357

<211> 43

<212> PRT

<213> Homo sapiens

<400> 357

Leu Glu Lys Ile Glu Ser Gly Glu Met Gly Leu Asn Lys Val Trp Ile

167

1 5 10 15
 Asn Ile Thr Glu Met Asn Lys Gln Ile Ser Leu Leu Thr Ser Ala Val
 20 25 30

Asn His Leu Lys Ala Asn Val Lys Ser Ala Ala
 35 40

<210> 358
 <211> 43
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (21)
 <223> Xaa equals any of the naturally occurring L-amino acids.

<400> 358
 Asp Leu Ile Ser Leu Pro Thr Thr Val Glu Gly Leu Gln Lys Ser Val
 1 5 10 15

Ala Ser Ile Gly Xaa Thr Leu Asn Ser Val His Leu Ala Val Glu Ala
 20 25 30

Leu Gln Lys Thr Val Asp Glu His Lys Lys Thr
 35 40

<210> 359
 <211> 43
 <212> PRT
 <213> Homo sapiens

<400> 359
 Met Glu Leu Leu Gln Ser Asp Met Asn Gln His Phe Leu Lys Glu Thr
 1 5 10 15

Pro Gly Ser Asn Gln Ile Ile Pro Ser Pro Ser Ala Thr Ser Glu Leu
 20 25 30

Asp Asn Lys Thr His Ser Glu Asn Leu Lys Gln
 35 40

<210> 360
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 360
 Met Gly Asp Arg Ser Ala Thr Leu Lys Arg Gln Ser Leu Asp Gln Val
 1 5 10 15

Thr Asn Arg Thr Asp Thr Val Lys Ile Gln Ser Ile Lys Lys Glu Gly
 20 25 30

<210> 361
 <211> 12

168

<212> PRT

<213> Homo sapiens

<400> 361

Ser Pro Gln Phe Leu Ser Ser Lys Ser Leu Pro Thr
 1 5 10

<210> 362

<211> 107

<212> PRT

<213> Homo sapiens

<400> 362

Gly Pro Pro Ser Pro Arg Gly Leu Pro Ser Leu Pro Leu His Leu Pro
 1 5 10 15

Ala Pro Arg Arg Tyr Leu Gln Ser Arg Tyr Ala Cys Ser Gln Ser Ser
 20 25 30

Val Ser Ala Ala Ala Arg Arg Trp Gly Ser Gly Trp Met Ala Trp Asp
 35 40 45

Pro Trp Asn Gln Ala Ser Gly Arg Tyr Ala Arg Ile Thr Leu Leu Ser
 50 55 60

Val Gln Ala Cys His Gln Pro Thr Val Trp Pro Arg Ala Gly His Ser
 65 70 75 80

Leu Pro Glu Arg Tyr Ser Leu His Pro His Asn Gly Asp Ser Thr His
 85 90 95

Leu Ser Gly Leu Leu Thr Val Lys Cys Gly Ala
 100 105

<210> 363

<211> 37

<212> PRT

<213> Homo sapiens

<400> 363

Gly Pro Pro Ser Pro Arg Gly Leu Pro Ser Leu Pro Leu His Leu Pro
 1 5 10 15

Ala Pro Arg Arg Tyr Leu Gln Ser Arg Tyr Ala Cys Ser Gln Ser Ser
 20 25 30

Val Ser Ala Ala Ala
 35

<210> 364

<211> 33

<212> PRT

<213> Homo sapiens

<400> 364

Arg Arg Trp Gly Ser Gly Trp Met Ala Trp Asp Pro Trp Asn Gln Ala
 1 5 10 15

Ser Gly Arg Tyr Ala Arg Ile Thr Leu Leu Ser Val Gln Ala Cys His
 20 25 30

169

Gln

<210> 365
<211> 37
<212> PRT
<213> Homo sapiens

<400> 365
Pro Thr Val Trp Pro Arg Ala Gly His Ser Leu Pro Glu Arg Tyr Ser
1 5 10 15
Leu His Pro His Asn Gly Asp Ser Thr His Leu Ser Gly Leu Leu Thr
20 25 30
Val Lys Cys Gly Ala
35

<210> 366
<211> 18
<212> PRT
<213> Homo sapiens

<400> 366
Asn Gln Glu Asn Ser Leu Gln Thr Asn Ser Tyr Leu Asp Ser Thr Glu
1 5 10 15
Ser Lys

<210> 367
<211> 31
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (17)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (19)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (30)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 367
Gln Lys Arg Ala Cys Phe Pro Phe Ala Phe Cys Arg Asp Cys Gln Phe
1 5 10 15
Xaa Glu Xaa Ser Pro Ala Met Leu Pro Val Gln Pro Ala Xaa Leu
20 25 30

<210> 368
<211> 11
<212> PRT

170

<213> Homo sapiens

<400> 368

Val Ser Ala His Gly Ile Trp Leu Phe Arg Ser
1 5 10

<210> 369

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (35)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (37)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (48)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 369

Lys His Ala Ala Pro Pro Ala Ser Leu Ser Leu Ser Leu Leu Leu His
1 5 10 15His Gly Gln Lys Arg Ala Cys Phe Pro Phe Ala Phe Cys Arg Asp Cys
20 25 30Gln Phe Xaa Glu Xaa Ser Pro Ala Met Leu Pro Val Gln Pro Ala Xaa
35 40 45

Leu

<210> 370

<211> 12

<212> PRT

<213> Homo sapiens

<400> 370

Ile Pro Glu Glu Ala Ser Cys Phe Pro Ser Ala Val
1 5 10

<210> 371

<211> 17

<212> PRT

<213> Homo sapiens

<400> 371

Glu Ile Leu Phe Gly Lys Leu Lys Ser Lys Ala Ala Leu Cys Thr Gln
1 5 10 15

Gly

171

<210> 372
<211> 19
<212> PRT
<213> Homo sapiens

<400> 372
His Ala Asp Arg Tyr Thr Cys Cys Arg Cys Leu Ser Pro Phe Ser Leu
1 5 10 15
Ala Gly Leu

<210> 373
<211> 15
<212> PRT
<213> Homo sapiens

<400> 373
Leu Ser Asp Pro Leu Leu Leu Pro Asp Cys Ser Phe Ser Phe Asn
1 5 10 15

<210> 374
<211> 25
<212> PRT
<213> Homo sapiens

<400> 374
Lys Ala Val Ala Tyr Ala Asn Val Ser Cys Arg Arg Phe Lys His Lys
1 5 10 15
Thr Thr Lys Leu Gly Pro Ile Gln Trp
20 25

<210> 375
<211> 26
<212> PRT
<213> Homo sapiens

<400> 375
Pro Ser Ser Gln Ser Pro Glu Pro Pro Gln Pro Leu Ser Leu Phe Val
1 5 10 15
Thr Arg Leu Pro Asn Leu Tyr Asp Phe Pro
20 25

<210> 376
<211> 19
<212> PRT
<213> Homo sapiens

<400> 376
Ser Arg Gln Ile Ile Cys Thr Asn Leu Cys Lys Cys Thr Pro Ile Cys
1 5 10 15
Phe Leu Phe

<210> 377
<211> 11

172

<212> PRT
<213> Homo sapiens

<400> 377
Met Leu Leu Pro Val Asn Thr Leu Leu Tyr Ile
1 5 10

<210> 378
<211> 14
<212> PRT
<213> Homo sapiens

<400> 378
Leu Leu Thr Pro Leu Cys Phe Phe Tyr Gly Thr Ser Arg Pro
1 5 10

<210> 379
<211> 7
<212> PRT
<213> Homo sapiens

<400> 379
Pro Tyr Leu Glu Leu Val Thr
1 5

<210> 380
<211> 13
<212> PRT
<213> Homo sapiens

<400> 380
Leu Leu Lys Lys Lys Lys Gln Ser Val Gly Phe Ser Val
1 5 10

<210> 381
<211> 7
<212> PRT
<213> Homo sapiens

<400> 381
Cys Ile Leu Glu Ala Gly Arg
1 5

<210> 382
<211> 11
<212> PRT
<213> Homo sapiens

<400> 382
Met Gly Phe Ser Ala Pro Thr Pro Gly Pro Leu
1 5 10

<210> 383
<211> 11
<212> PRT
<213> Homo sapiens

<400> 383

173

Phe Asp Leu Arg Arg Leu Ile Leu Ser Ile Val
1 5 10

<210> 384
<211> 17
<212> PRT
<213> Homo sapiens

<400> 384
Ala Phe Cys Pro His Val Thr Pro Cys Lys Tyr Ala Val Ile His Thr
1 5 10 15

Val

<210> 385
<211> 11
<212> PRT
<213> Homo sapiens

<400> 385
Asn Thr Pro Leu Leu Phe Leu Trp Asp Leu Gln
1 5 10

<210> 386
<211> 17
<212> PRT
<213> Homo sapiens

<400> 386
Ala Thr Ile Phe Arg Thr Ser Tyr Leu Ile Lys Lys Glu Lys Thr Val
1 5 10 15

Cys

<210> 387
<211> 17
<212> PRT
<213> Homo sapiens

<400> 387
Trp Leu Leu Ser Leu His Leu Gly Gly Arg Glu Val Arg Ala Gly Ala
1 5 10 15

Pro

<210> 388
<211> 11
<212> PRT
<213> Homo sapiens

<400> 388
Gln Thr Leu Gln Glu Gly Ser Leu His Ser Ile
1 5 10

<210> 389

<211> 95
 <212> PRT
 <213> Homo sapiens

<400> 389

Met Gly Phe Ser Ala Pro Thr Pro Gly Pro Leu Phe Asp Leu Arg Arg
 1 5 10 15
 Leu Ile Leu Ser Ile Val Ala Phe Cys Pro His Val Thr Pro Cys Lys
 20 25 30
 Tyr Ala Val Ile His Thr Val Asn Thr Pro Leu Leu Phe Leu Trp Asp
 35 40 45
 Leu Gln Ala Thr Ile Phe Arg Thr Ser Tyr Leu Ile Lys Lys Glu Lys
 50 55 60
 Thr Val Cys Trp Leu Leu Ser Leu His Leu Gly Arg Glu Val Arg
 65 70 75 80
 Ala Gly Ala Pro Gln Thr Leu Gln Glu Gly Ser Leu His Ser Ile
 85 90 95

<210> 390
 <211> 33
 <212> PRT
 <213> Homo sapiens

<400> 390

Tyr Trp Val Ser Ile Ser Gln Arg Ser Val Cys Gln Gln Ala Arg Thr
 1 5 10 15
 Ser Ile Phe Phe Lys Asp Gly Leu Ser Arg Glu Lys Tyr Ser Asn Asn
 20 25 30

Gly

<210> 391
 <211> 160
 <212> PRT
 <213> Homo sapiens

<400> 391

Leu Ser Val Arg Ala Pro Gly Val Pro Ala Ala Arg Pro Arg Leu Ser
 1 5 10 15
 Ser Ala Arg Gln Ala Gly Ala Gly Arg Gly Glu Leu Arg Gly Gln Arg
 20 25 30
 Leu Trp Leu Gly Pro Gln Cys Gly Cys Gly Ala Gly Gln Ala Gly Ser
 35 40 45
 Met Leu Arg Ala Val Gly Ser Leu Leu Arg Leu Gly Arg Gly Leu Thr
 50 55 60
 Val Arg Cys Gly Pro Gly Ala Pro Leu Glu Ala Thr Arg Arg Pro Ala
 65 70 75 80
 Pro Ala Leu Pro Pro Arg Gly Leu Pro Cys Tyr Ser Ser Gly Gly Ala
 85 90 95

175

Pro Ser Asn Ser Gly Pro Gln Gly His Gly Glu Ile His Arg Val Pro
 100 105 110

Thr Gln Arg Arg Pro Ser Gln Phe Asp Lys Lys Ile Leu Leu Trp Thr
 115 120 125

Gly Arg Phe Lys Ser Met Glu Glu Ile Pro Pro Arg Ile Pro Pro Glu
 130 135 140

Met Ile Asp Thr Ala Arg Asn Lys Ala Arg Val Lys Ala Cys Tyr Ile
 145 150 155 160

<210> 392
 <211> 36
 <212> PRT
 <213> Homo sapiens

<400> 392
 Leu Ser Val Arg Ala Pro Gly Val Pro Ala Ala Arg Pro Arg Leu Ser
 1 5 10 15

Ser Ala Arg Gln Ala Gly Ala Gly Arg Gly Glu Leu Arg Gly Gln Arg
 20 25 30

Leu Trp Leu Gly
 35

<210> 393
 <211> 34
 <212> PRT
 <213> Homo sapiens

<400> 393
 Pro Glu Cys Gly Cys Gly Ala Gly Gln Ala Gly Ser Met Leu Arg Ala
 1 5 10 15

Val Gly Ser Leu Leu Arg Leu Gly Arg Gly Leu Thr Val Arg Cys Gly
 20 25 30

Pro Gly

<210> 394
 <211> 34
 <212> PRT
 <213> Homo sapiens

<400> 394
 Ala Pro Leu Glu Ala Thr Arg Arg Pro Ala Pro Ala Leu Pro Pro Arg
 1 5 10 15

Gly Leu Pro Cys Tyr Ser Ser Gly Gly Ala Pro Ser Asn Ser Gly Pro
 20 25 30

Gln Gly

```

1              5              10              15
Thr Tyr

<210> 406
<211> 18
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (9)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 406
Pro Gly Gln Trp Val Arg Glu Ile Xaa Leu Val Gly Arg Ala Val Ala
 1              5              10              15
Arg Val

<210> 407
<211> 16
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 407
Leu Thr Trp Pro Pro Xaa Gly Pro Met Gly Thr Val Trp Pro Gly Phe
 1              5              10              15

<210> 408
<211> 17
<212> PRT
<213> Homo sapiens

<400> 408
Met Ala Asp Ile Pro Gly Thr Phe Leu Ala Leu Gly Cys His Gly Gln
 1              5              10              15
Arg

<210> 409
<211> 15
<212> PRT
<213> Homo sapiens

<400> 409
Val Gly Arg Gly Ser Trp Ala Ser Gly Trp Thr Asn Gln Ser Ala
 1              5              10              15

```

179

<210> 410
 <211> 16
 <212> PRT
 <213> Homo sapiens

<400> 410
 Pro Asp His Pro Leu Pro Val Gly Leu Leu Glu Ala Trp Arg Val Glu
 1 5 10 15

<210> 411
 <211> 94
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (84)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 411
 Leu Ala Arg Ala Asp Pro Pro Gly Cys Arg Arg Arg Gly Trp Arg Pro
 1 5 10 15

Ser Ser Ala Glu Leu Gln Leu Arg Leu Leu Thr Pro Thr Phe Glu Gly
 20 25 30

Ile Asn Gly Leu Leu Leu Lys Gln His Leu Val Gln Asn Pro Val Arg
 35 40 45

Leu Trp Gln Leu Leu Gly Gly Thr Phe Tyr Phe Asn Thr Ser Arg Leu
 50 55 60

Lys Gln Lys Asn Lys Glu Lys Asp Lys Ser Lys Gly Lys Ala Pro Glu
 65 70 75 80

Glu Asp Glu Xaa Glu Arg Arg Arg Arg Glu Arg Asp Asp Gln
 85 90

<210> 412
 <211> 12
 <212> PRT
 <213> Homo sapiens

<400> 412
 Phe Leu Arg Phe Trp Cys Thr Cys His Val Ser Ser
 1 5 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/09847

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) : C12N 15/00, 15/12; C07K 14/00, 14/435 US CL : 435/69.1, 320.1, 325; 536/23.5; 530/350 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/69.1, 320.1, 325; 536/23.5; 530/350 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Genbank, Genbak-EST, Swissprot		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KUSTIKOVA et al. Cloning of the tag7 gene expressed in metastatic mouse tumors. Genetika. May 1996, Vol. 32, No. 5, pages 621-628, Figure 3.	1-2, 7-11, 14-16
X,P	KISELEV et al. Molecular cloning and characterization of the mouse tag7 gene encoding a novel cytokine. J. Biol. Chem. July 1998, Vol. 273, No. 9, pages 18633-18639, Figure 1.	1-2, 7-11, 14-16
X,P	KANG et al. A peptidoglycan recognition protein in innate immunity conserved from insects to humans. Proc. Natl. Acad. Sci. USA. August 1998, Vol. 95, pages 10078-10082, Figure 1.	1-2, 7-11, 14-16
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents.	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family	
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 16 AUGUST 1999	Date of mailing of the international search report 10 SEP 1999	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer <i>D. Laurence</i> ELIANE LAZAR-WESLEY Telephone No. (703) 308-0196	

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/09847

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database on GENESEQ, Accession Number W23723, SELSTED ME., Murine granulocyte peptide A precursor (antimicrobial MGP-A), 18 February 1998.	1-2, 7-11, 14-16
X,P	Database on GENBANK, Accession Number AC002559, EVANS et al. Homo sapiens chromosome 10 PAC clone pDJ205g22. 13 November 1998.	1-2, 7-11, 14-16
X	Database on GENBANK, Accession Number AA594742, NCI-CGAP . no03g09sl Homo sapiens cDNA clone.	1-2,7-11, 14-16

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/09847

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-12, 14-16 and 21, as they apply to SEQ ID No:11 and SEQ ID No:111

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/09847

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-12, 14-16 and 21, drawn to a polynucleotide fragment of SEQ ID No:11, a polynucleotide encoding a polypeptide fragment of SEQ ID No:111, a vector, a method of making a host cell, a host cell, a polypeptide, a method of making a polypeptide, and a gene.

Group II, claim 13, drawn to an antibody.

Group III, claim 17, drawn to a method of treating or preventing a medical condition.

Group IV, claims 18-19, drawn to a method of diagnosis.

Group V, claim 20, drawn to a method of identifying a binding partner.

Group VI, claim 22, drawn to a method of identifying an activity in a biological assay.

Group VII, claim 23, drawn to a product.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: pursuant to 37 CFR 1.475(d), this Authority considers that the main invention in the instant application comprises the first recited product, which is the polynucleotide fragment of SEQ ID No:11, the polynucleotide encoding the polypeptide of SEQ ID No:111, and the first recited method of using that product, namely in the process of making a recombinant cell. Note that there is no method of making the polynucleotide. Also included in this group is the product made, namely the encoded polypeptide, and vector, host cell, method of making the polypeptide, and gene. Further, pursuant to 37 CFR 1.475(b)-(d), the ISA/US considers that the materially and functionally dissimilar products of Group II and VII, and the additional methods of Groups III-VI do not correspond to the main invention. This Authority therefore considers that the several inventions do not share a special technical feature within the meaning of PCT Rule 13.2 and thus do not relate to a single general inventive concept within the meaning of PCT rule 13.1.

